(Article begins on next page)

Gray Coding for Multilevel Constellations in Gaussian Noise

Erik Agrell, Johan Lassing, Erik G. Ström, and Tony Ottosson

***Abstract*— The problem of finding the optimal labeling (bit-to-symbol mapping) of multilevel coherent PSK, PAM, and QAM constellations with respect to minimizing the bit error probability (BEP) over a Gaussian channel is addressed. We show that using the binary reflected Gray code (BRGC) to label the signal constellation results in the lowest possible BEP for high enough signal energy-to-noise ratios and analyze what is "high enough" in this sense. It turns out that the BRGC is optimal for PSK and PAM systems whenever the target BEP is at most a few percent, which covers most systems of practical interest. New and simple closed-form expressions are presented for the BEP of PSK, PAM, and QAM using the BRGC.**

***Index Terms*— binary reflected Gray code, bit error probability, bit error rate, constellation labeling, digital modulation, Gray mapping, optimal labeling, PAM, PSK, QAM.**

I. INTRODUCTION

This paper addresses the problem of selecting an optimal labeling with respect to minimizing the bit error probability (BEP) in digital communication systems with coherent symbol detection over an additive white Gaussian noise (AWGN) channel [1]. We study transmission of equally likely, statistically independent bits using multilevel phase shift keying (PSK), pulse amplitude modulation (PAM), and quadrature amplitude modulation (QAM) systems; with the binary reflected Gray code (BRGC), other Gray codes, and other labelings; for finite and infinite signal-to-noise ratios (SNR). Only uncoded transmission (or more precisely, coding without redundancy) is considered. The corresponding problem in systems with error-correcting codes is considered in, e.g., [2].

It is established engineering knowledge that labeling signal constellations with Gray codes (in particular, the BRGC [3]) is a way to reduce the BEP for the systems considered in this paper. There exist, however, a multitude of nonequivalent Gray codes. The theoretical question whether the BRGC is the best way to label the constellations has so far been an open question even in the asymptotic case of infinite SNR, although this is sometimes stated as a fact in the literature [4]. That the use of the BRGC (or even Gray codes) is not optimal for all SNR's for at least some constellation sizes can be demonstrated by explicit evaluation of the BEP for various labelings and modulation forms (see [5, Fig. 8] for a 64-QAM example).

Manuscript submitted June 29, 2005; revised May 30, 2006. The material in this paper was published in part in Tech. Rep. R001/2005, Dept. of Signals and Systems, Chalmers University of Technology, Sweden, Jan. 2005.

E. Agrell, E. G. Ström, and T. Ottosson are with the Communication Systems group, Department of Signals and Systems, Chalmers University of Technology, 41296 Göteborg, Sweden (e-mail: agrell@chalmers.se; erik.strom@chalmers.se; tony.ottosson@chalmers.se). J. Lassing is with Qamcom Technology, Gruvgatan 8, 42130 Västra Frölunda, Sweden (e-mail: johan.lassing@qamcom.se).

The BEP of the systems considered herein was shown in [5] to be a function of two quantities; the average distance spectrum (ADS), derived from the constellation labeling, and the communication channel. We established in [5] the somewhat artificial result that the BRGC is the optimum labeling for PSK and PAM with respect to certain properties of the ADS, and the question whether the BRGC also yields minimum BEP over a practical channel with a finite (and in some cases even for infinite) SNR was left unanswered. In this paper, the optimality criterion is therefore changed from the one in [5] to the more relevant requirement that the optimal labeling should minimize the BEP of the communication system. We assume an AWGN channel and show that the minimum achievable BEP is, indeed, obtained by using the BRGC as long as the signal energy-to-noise ratio is higher than a finite threshold, which depends on the modulation scheme and the size of the constellation.

The paper is organized as follows. In Section II, the preliminaries are presented and a proof method is outlined for the optimality of the BRGC. Section III presents general BEP expressions for each of the three studied modulation formats, which hold for arbitrary labelings. In Section IV, we derive a particularly useful partitioning of the set of all possible labelings (competing with the BRGC on being the optimal labeling). This partitioning is used in Sections V–VII to prove the optimality of the BRGC for PSK, PAM, and QAM, respectively. In Section VIII, the analysis from Section III is continued and specialized to the case of BRGC labelings, resulting in explicit, closed-form expressions for the ADS and BEP of the BRGC. Finally, conclusions and comments are given in Section IX.

II. PRELIMINARIES

Some of the most central definitions and notation for the paper are collected in this section. The proof method is also outlined and two important analytical results, central to the proofs, are given.

A. Definitions and Notation

The presented work deals with binary labelings and in this subsection we introduce the nomenclature and definitions that are used in the discussion.

A *binary labeling* λ of order $m \in \mathbb{Z}^+$ is defined as a sequence of $M = 2^m$ distinct vectors (labels or codewords), $\lambda = (\boldsymbol{c}_0, \boldsymbol{c}_1, \ldots, \boldsymbol{c}_{M-1})$, where each $\boldsymbol{c}_i \in \{0,1\}^m$. A *rectangular binary labeling* λ of order $(m_1, m_2) \in \mathbb{Z}^+ \times \mathbb{Z}^+$ consists of all vectors (labels) in $\{0,1\}^{m_1+m_2}$, arranged in a matrix of dimension $M_1 = 2^{m_1}$ by $M_2 = 2^{m_2}$.

TABLE I
THE BINARY REFLECTED GRAY CODES OF ORDERS $m = 1, 2, 3,$ AND 4.

β_1	β_2	β_3	β_4
0	0 0	0 0 0	0 0 0 0
1	0 1	0 0 1	0 0 0 1
	1 1	0 1 1	0 0 1 1
	1 0	0 1 0	0 0 1 0
		1 1 0	0 1 1 0
		1 1 1	0 1 1 1
		1 0 1	0 1 0 1
		1 0 0	0 1 0 0
			1 1 0 0
			1 1 0 1
			1 1 1 1
			1 1 1 0
			1 0 1 0
			1 0 1 1
			1 0 0 1
			1 0 0 0

A binary labeling λ' of order m is said to be *optimal* for signal energy-to-noise ratio γ if

$$P_{\mathrm{b}}(\lambda', \gamma) \leq P_{\mathrm{b}}(\lambda, \gamma)$$

for all labelings λ of order m, where $P_{\mathrm{b}}(\lambda, \gamma)$ is the bit error probability, which is defined for PSK, PAM, and QAM systems in Section III, and γ is defined as E_{s}/N_0, where E_{s} is the symbol energy and $N_0/2$ is the two-sided power spectral density of the AWGN.

Throughout the paper, we discuss a particular class of labelings; Gray codes. A *binary Gray code* of order m is a binary labeling with $M = 2^m$ distinct labels, where adjacent labels differ in only one of the m positions. If we impose the additional requirement that the first and the last labels differ in a single position, the labeling is said to be a *cyclic binary Gray code*. Analogously, a *rectangular binary Gray code* is a rectangular labeling where adjacent labels, horizontally as well as vertically, differ in only one bit.

Among the cyclic Gray codes, we are especially interested in the *binary reflected Gray code* (BRGC) [3, 5], and we denote the BRGC of order m by β_m. For reference, we have listed β_m for $m = 1, \ldots, 4$ in Table I. With a *two-dimensional BRGC* we mean the direct product of two BRGC's. The BRGC can be transformed into a number of labelings that yield exactly the same BEP by means of some trivial operations such as interchanging bit positions. The role of such operations was discussed in some depth in [5], and we will not say more about them here. We loosely refer to the class of all such labelings as *the* BRGC.

The ADS of the labeling λ, denoted by $\bar{d}(k, \lambda)$ for PSK, essentially tells us the average number of bits that differ in two labels separated by k symbols in the constellation. The precise definitions depend on the constellation and follow in Section III. Letting Λ_m denote the set of all labelings having an ADS that differs from the ADS of β_m, we define the *critical index* $T(\lambda)$ of a PSK labeling $\lambda \in \Lambda_m$ as

$$T(\lambda) \triangleq \min\{k \in \mathbb{Z}^+ : \bar{d}(k, \lambda) \neq \bar{d}(k, \beta_m)\}. \tag{1}$$

From [5, Th. 5] we know that

$$\bar{d}(k, \lambda) > \bar{d}(k, \beta_m) \tag{2}$$

at $k = T(\lambda)$. The definition (1) also applies to PAM if $\bar{d}(k, \lambda)$ is replaced with $\bar{h}(k, \lambda)$ (defined in Section III-B).

The set of all critical indices is the *critical index set*

$$\Psi_m \triangleq \{T(\lambda) : \lambda \in \Lambda_m\}. \tag{3}$$

In addition, it will be convenient to have a designator for the set of labelings for which $T(\lambda) = i$,

$$\Lambda_m(i) \triangleq \{\lambda \in \Lambda_m : T(\lambda) = i\}. \tag{4}$$

Although obvious from (3) and (4), we explicitly state that

$$\Lambda_m = \bigcup_{i \in \Psi_m} \Lambda_m(i)$$

and

$$\Lambda_m(i) \cap \Lambda_m(j) = \varnothing \qquad \text{for } i \neq j$$

since these relations are central to the proof method as described in the next subsection.

B. Outline of Proof Method

Before proceeding to the details, we give an outline of the proof method that will be used. Using the definitions and the notation introduced in the previous subsection, the aim of this paper is for each m and each modulation form to establish a range of γ for which

$$P_{\mathrm{b}}(\beta_m, \gamma) \leq \min_{\lambda \in \Lambda_m} P_{\mathrm{b}}(\lambda, \gamma) \tag{5}$$

that is, for what signal energy-to-noise ratios the labeling β_m will result in the lowest BEP among all possible labelings. We define the *optimality threshold* γ_m^* for order m as the smallest value such that (5) holds for all $\gamma \geq \gamma_m^*$.

We will address (5) by using the equivalent formulation

$$0 \leq \min_{i \in \Psi_m} \left[\min_{\lambda \in \Lambda_m(i)} P_{\mathrm{b}}(\lambda, \gamma) - P_{\mathrm{b}}(\beta_m, \gamma) \right]. \tag{6}$$

We will, for each $i \in \Psi_m$, lowerbound the expression inside the brackets in (6) and establish a range of γ for which the bound is non-negative. This yields an upper bound $\hat{\gamma}_m$ on γ_m^*, which is computed separately for PSK, PAM, and QAM.

C. Two Lemmas of Monotonicity

In order to find the range of γ for which (5) is valid, we make use of two results from calculus, which are derived in this subsection.

Lemma 1: For constants a and b such that $0 \leq a < b$, consider the difference $\Delta(x) = f(bx) - f(ax)$. If

- $f(x)$ is continuous and twice differentiable for $x \geq 0$,

- $f'(x) > 0$ for $x > 0$, and
- $f''(x) > 0$ for $x > 0$,

then $\Delta(x)$ is a strictly increasing function in x for $x \geq 0$.

Proof: Since $f''(x) > 0$ for $x > 0$, we have for $0 \leq a < b$ and $x > 0$

$$0 < \int_{ax}^{bx} f''(t)\,dt = f'(bx) - f'(ax).$$

Since $f'(x) \geq 0$ for $x > 0$, we have

$$f'(bx) - f'(ax) \leq f'(bx) - \frac{a}{b} f'(ax) = \frac{\Delta'(x)}{b}$$

for $0 \leq a < b$, showing that $\Delta'(x) > 0$, which completes the proof. □

The next lemma involves the Gaussian Q-function

$$Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2}\,dt.$$

Lemma 2: For two constants a and b, such that $0 \leq a < b$, the ratio

$$r(x) = \frac{Q(ax)}{Q(bx)}$$

is a strictly increasing function of x for $x > 0$.

Proof: Let $f(x) = -\log Q(x)$, which is a continuous, twice differentiable function for all x with first derivative

$$f'(x) = -\frac{Q'(x)}{Q(x)}.$$

For the second derivative we have

$$f''(x) = \frac{Q'(x)^2 - Q''(x)Q(x)}{Q(x)^2}$$

and since

$$Q'(x) = -\frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} \quad \text{and} \quad Q''(x) = \frac{xe^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}$$

we have $f'(x) > 0$ for all x and

$$f''(x) = \frac{xe^{-\frac{1}{2}x^2}}{\sqrt{2\pi}Q(x)^2}\left[\frac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}} - Q(x)\right].$$

Now, as $e^{-\frac{1}{2}x^2}/x\sqrt{2\pi}$ is a well known upper bound on $Q(x)$ for $x > 0$ [6, p. 98], we conclude that for $x > 0$, we have $f''(x) > 0$. Applying Lemma 1 to $f(x)$, we find that

$$\Delta(x) = -\log Q(bx) + \log Q(ax) = \log \frac{Q(ax)}{Q(bx)}$$

is a strictly increasing function for $x > 0$ and $0 \leq a < b$, which also implies that $r(x) = e^{\Delta(x)}$ is a strictly increasing function of x for $x > 0$. □

III. BEP of Systems with Any Labeling

This section provides simple, closed-form expressions for the BEP of each of the three studied modulation formats. They all separate the influence of the channel from that of the labeling, where the latter is captured by the ADS.

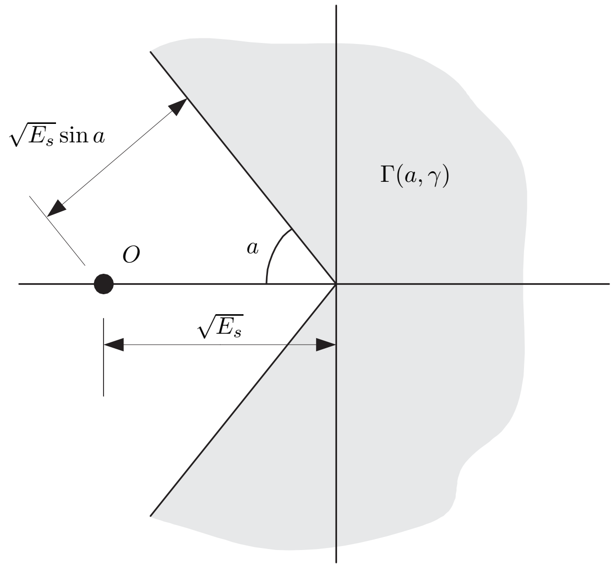

Fig. 1. The shaded area represents the probability $\Gamma(a,\gamma)$ that would result from integration of a Gaussian pdf with variance $N_0/2$ in each dimension, centered on O, over this region.

A. Bit Error Probability for PSK

The average BEP of M-PSK, where $M = 2^m$ for any integer $m \geq 1$, over AWGN channels can be written [7]

$$P_{\mathrm{b}}(\lambda,\gamma) = \frac{1}{m}\sum_{k=1}^{M-1} \bar{d}(k,\lambda)P(k,\gamma) \tag{7}$$

where $\bar{d}(k,\lambda)$ is the ADS of an M-PSK constellation labeling λ. It is defined for all integers k as

$$\bar{d}(k,\lambda) \triangleq \frac{1}{M}\sum_{l=0}^{M-1} d_{\mathrm{H}}\left(\boldsymbol{c}_l, \boldsymbol{c}_{(l+k) \bmod M}\right) \tag{8}$$

where $\boldsymbol{c}_k$ is the m-bit binary label assigned to the kth constellation point and the Hamming distance $d_{\mathrm{H}}(\boldsymbol{c}_j, \boldsymbol{c}_k)$ is the number of positions in which $\boldsymbol{c}_j$ and $\boldsymbol{c}_k$ differ. The ADS denotes the average number of bits that differ between binary labels assigned to constellation points separated by k steps in the PSK constellation. If it is clear from the context which labeling λ is concerned, we will simply write $\bar{d}(k)$ for the ADS. The *crossover probability* $P(k,\gamma)$ is the probability that the received signal vector is found in a decision region belonging to a signal point k steps away (clockwise along the PSK circle) from the transmitted signal point.

To find an expression for $P(k,\gamma)$ for a given symbol energy-to-noise ratio $\gamma \triangleq E_{\mathrm{s}}/N_0$, we refer to Figure 1 and consider a rotationally invariant, two-dimensional Gaussian probability density function (pdf) with variance $N_0/2$ per dimension, centered on the point O. In the two-dimensional setting considered herein, the noncentral t-distribution gives the probability $\Gamma(a,\gamma)$, which for $0 \leq a \leq \pi$ denotes the integral of the Gaussian pdf over the region bounded by angles $\pm a$ not containing O. For $k = 1,\ldots,M/2-1$, the probability $P(k,\gamma)$ is related to $\Gamma(a,\gamma)$ through the relation

$$P(k,\gamma) = \frac{1}{2}\left[\Gamma\left(\frac{(2k-1)\pi}{M},\gamma\right) - \Gamma\left(\frac{(2k+1)\pi}{M},\gamma\right)\right] \triangleq \frac{1}{2}\left[\Gamma\left(a_k,\gamma\right) - \Gamma\left(b_k,\gamma\right)\right] \qquad (9)$$

while for $k = 0$,

$$P(0,\gamma) = 1 - \Gamma\left(\frac{\pi}{M},\gamma\right)$$

and for $k = M/2$, we have

$$P(M/2,\gamma) = \Gamma\left(\pi - \frac{\pi}{M},\gamma\right). \qquad (10)$$

By symmetry, $P(k,\gamma) = P(M-k,\gamma)$ for $k = M/2+1,\ldots, M-1$.

There exist several expressions for the probability $\Gamma(a,\gamma)$, which is closely related to the noncentral t-distribution [8, 9], for example

$$\Gamma(a,\gamma) = \frac{1}{\pi}\int_0^{\pi-a} e^{-\gamma\frac{\sin^2 a}{\sin^2\varphi}}\,d\varphi \qquad (11)$$

$$= 2Q(\sqrt{2\gamma}\sin a) - \frac{1}{\pi}\int_0^{a} e^{-\gamma\frac{\sin^2 a}{\sin^2\varphi}}\,d\varphi. \qquad (12)$$

For numerical stability, we prefer (11) if $a \geq \pi/2$ and (12) otherwise. The expression (11) was given in [10, p. 198] and (12) can be proved using $\Gamma(\pi/2,\gamma) = Q(\sqrt{2\gamma})$, see Figure 1.

We may simplify the above BEP expressions further. By inserting (9) and (10) in (7), we find that

$$P_{\mathrm{b}}(\lambda,\gamma) = \frac{1}{m}\sum_{k=1}^{M/2}\bar{\Delta}(k,\lambda)\Gamma\left(a_k,\gamma\right) \qquad (13)$$

where

$$\bar{\Delta}(k,\lambda) \triangleq \bar{d}(k,\lambda) - \bar{d}(k-1,\lambda) \qquad (14)$$

is the *differential ADS*.

B. Bit Error Probability for PAM

The BEP expression for M-PAM can be written in a form similar to (7) and is again a function of the labeling λ used to label the constellation and the signal energy-to-noise ratio γ [5]

$$P_{\mathrm{b}}(\lambda,\gamma) = \frac{2}{m}\sum_{k=1}^{\infty}\bar{h}(k,\lambda)\mathcal{P}(k,\mu(\gamma)) \qquad (15)$$

where

$$\mu(\gamma) = \sqrt{\frac{6\gamma}{M^2-1}} \qquad (16)$$

is half the distance between adjacent constellation vectors divided by $\sqrt{N_0/2}$. In the rest of this paper, $\mu(\gamma)$ will sometimes be written as μ, letting the dependence on γ be implicit. Furthermore, $\mathcal{P}(k,\mu)$ is expressed in terms of the Gaussian Q-function as

$$\mathcal{P}(k,\mu) \triangleq Q\left((2k-1)\mu\right) - Q\left((2k+1)\mu\right). \qquad (17)$$

The ADS $\bar{h}(k,u)$ of any sequence $u = (\boldsymbol{c}_0,\ldots,\boldsymbol{c}_{M-1})$ of M binary vectors is defined for all integers k as

$$\bar{h}(k,u) \triangleq \frac{1}{2M}\sum_{l=0}^{M-1}\left(d_{\mathrm{H}}(\boldsymbol{c}'_l,\boldsymbol{c}'_{l+k}) + d_{\mathrm{H}}(\boldsymbol{c}'_l,\boldsymbol{c}'_{l-k})\right) \qquad (18)$$

with $\boldsymbol{c}'_i \triangleq \boldsymbol{c}_{r(M,i)}$, where r is a ramp function given by

$$r(M,i) \triangleq \begin{cases} 0, & i < 0 \\ i, & 0 \leq i \leq M-1 \\ M-1, & i > M-1. \end{cases} \qquad (19)$$

As for the PSK case, we will write $\bar{h}(k)$ for the ADS if it is obvious from the context which sequence u is concerned.

It follows straightforwardly from this definition that for any sequence u, $\bar{h}(0,u) = 0$. More importantly, for the special case when $u = \lambda$ is a labeling, we note that for $k \geq M-1$, (18) counts the average number of ones per label taken over the entire labeling. For any labeling this average is $m/2$, so that $\bar{h}(k,\lambda) = m/2$ for $k \geq M-1$. This fact can be exploited to reduce the number of terms in (15) and obtain

$$P_{\mathrm{b}}(\lambda,\gamma) = Q\left((2M-3)\mu\right) + \frac{2}{m}\sum_{k=1}^{M-2}\bar{h}(k,\lambda) \cdot \left[Q\left((2k-1)\mu\right) - Q\left((2k+1)\mu\right)\right]. \qquad (20)$$

As in the case of M-PSK, we may simplify the expression further. Taking differences of $\bar{h}(k,\lambda)$ instead of the Q-functions, we obtain

$$P_{\mathrm{b}}(\lambda,\gamma) = \frac{2}{m}\sum_{k=1}^{M-1}\bar{\delta}(k,\lambda)Q\left((2k-1)\mu\right) \qquad (21)$$

where it follows from (18) and (19) that the differential ADS is

$$\bar{\delta}(k,\lambda) \triangleq \bar{h}(k,\lambda) - \bar{h}(k-1,\lambda) = \frac{1}{2M}\sum_{l=0}^{M-k-1}\big(2d_H(\boldsymbol{c}_l,\boldsymbol{c}_{l+k}) - d_H(\boldsymbol{c}_l,\boldsymbol{c}_{l+k-1}) - d_H(\boldsymbol{c}_{l+1},\boldsymbol{c}_{l+k})\big). \qquad (22)$$

The expressions (21)–(22) provide a convenient method to evaluate the BEP of PAM systems with any labeling.

C. Bit Error Probability for QAM

We consider rectangular $M_1 \times M_2$ QAM constellations such that $m_1 = \log_2 M_1$ and $m_2 = \log_2 M_2$ are integers, which are labeled by binary labels of length $m_1 + m_2$. To evaluate the BEP for QAM, we define virtual labels even for imaginary constellation vectors outside the $M_1 \times M_2$ grid as

$$\boldsymbol{c}'_{i,j} = \boldsymbol{c}_{r(M_1,i),r(M_2,j)}$$

for all integers i and j, where the ramp function r was defined in (19).

The BEP for this system, when used for transmission over an AWGN channel with average signal energy-to-noise ratio γ, can be written as

$$P_{\mathrm{b}}(\lambda,\gamma)=\frac{1}{m_1+m_2}\sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty}\bar{g}(k,l,\lambda)\mathcal{P}(k,\mu)\mathcal{P}(l,\mu)$$

where $\mathcal{P}$ is the same as in (17) and μ is still half the distance between adjacent constellation vectors divided by $\sqrt{N_0/2}$; the relation between μ and γ is for QAM (cf. (16))

$$\mu(\gamma)=\sqrt{\frac{6\gamma}{M_1^2+M_2^2-2}}. \tag{23}$$

The labeling λ is now rectangular (see Section II-A) and

$$\bar{g}(k,l,\lambda)\triangleq\frac{1}{M_1M_2}\sum_{i=0}^{M_1-1}\sum_{j=0}^{M_2-1}d_{\mathrm{H}}(\boldsymbol{c}'_{i,j},\boldsymbol{c}'_{i+k,j+l}).$$

To exploit the symmetry of the constellation, as was done for PAM in (15), we form an ADS with two components $\bar{t}$ and $\bar{t}_0$ by averaging components of $\bar{g}$ in groups of four. This yields

$$P_{\mathrm{b}}(\lambda,\gamma)=\frac{4}{m_1+m_2}\sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\bar{t}(k,l,\lambda)\mathcal{P}(k,\mu)\mathcal{P}(l,\mu)+\frac{4\mathcal{P}(0,\mu)}{m_1+m_2}\sum_{k=1}^{\infty}\bar{t}_0(k,\lambda)\mathcal{P}(k,\mu) \tag{24}$$

where for any integers k and l (dropping the dependence on λ to simplify the notation)

$$\bar{t}(k,l)\triangleq\frac{1}{4}\left[\bar{g}(k,l)+\bar{g}(k,-l)+\bar{g}(-k,l)+\bar{g}(-k,-l)\right] \tag{25}$$

$$\bar{t}_0(k)\triangleq\frac{1}{4}\left[\bar{g}(0,k)+\bar{g}(0,-k)+\bar{g}(k,0)+\bar{g}(-k,0)\right]. \tag{26}$$

The expression (24) is suitable for analytically comparing the performance of various labelings, as will be done in Section VII. To numerically evaluate the BEP of a given labeling, however, the infinite summations can be replaced by a finite number of terms, because

$$\bar{t}(k,l)=\bar{t}(\min\{k,M_1-1\},\min\{l,M_2-1\})$$
$$\bar{t}_0(k)=\bar{t}_0(\min\{k,\max\{M_1,M_2\}-1\}).$$

The calculations, which are not detailed here, follow in perfect analogy with (20)–(22).

So far, the expressions in this section hold for arbitrary QAM labelings. In the special case when the QAM labeling is the direct product of two PAM labelings, the expressions can be simplified further. Indeed, it can be shown that for any product labeling $\lambda_1\times\lambda_2$, the two-dimensional ADS components (25)–(26) get the particularly simple forms

$$\bar{t}(k,l)=\bar{h}(k,\lambda_1)+\bar{h}(l,\lambda_2)$$
$$\bar{t}_0(k)=\frac{\bar{h}(k,\lambda_1)+\bar{h}(k,\lambda_2)}{2}.$$

Substituting these expressions into (24) and simplifying yields the BEP [11]

$$P_{\mathrm{b}}(\lambda_1\times\lambda_2,\gamma)=\frac{m_1}{m_1+m_2}P_{\mathrm{b}}\left(\lambda_1,\frac{M_1^2-1}{M_1^2+M_2^2-2}\gamma\right)+\frac{m_2}{m_1+m_2}P_{\mathrm{b}}\left(\lambda_2,\frac{M_2^2-1}{M_1^2+M_2^2-2}\gamma\right) \tag{27}$$

where $P_{\mathrm{b}}(\lambda_1,\gamma)$ and $P_{\mathrm{b}}(\lambda_2,\gamma)$ are the BEP's of the constituent M_1-PAM and M_2-PAM systems, obtained from (21).

IV. The Critical Index Set

In order to address (6), we need to find the critical set Ψ_m defined in (3). We will rely on a method called labeling *expansion*, which is a way to construct a labeling λ_m of order m from a labeling λ_{m-1} of order $m-1$ [5]. For a labeling λ_{m-1} that is expanded into λ_m the following relations hold for $m\geq 2$ and $k\in\mathbb{Z}$

$$\bar{d}(4k,\lambda_m)=\bar{d}(2k,\lambda_{m-1})+f_1 \tag{28}$$
$$\bar{d}(4k+2,\lambda_m)=\bar{d}(2k+1,\lambda_{m-1})+f_2$$
$$\bar{d}(2k+1,\lambda_m)=\frac{\bar{d}(k,\lambda_{m-1})+\bar{d}(k+1,\lambda_{m-1})}{2}+f_3 \tag{29}$$

where f_1, f_2, and f_3 are functions of k and m, but independent of λ_{m-1}. The same relations, with different f_1, f_2 and f_3, hold for $\bar{h}(k,\lambda)$. An important property of labeling expansion is that expanding β_{m-1} gives β_m. Furthermore, we derive the following property of expanded cyclic labelings from [5, Lemma 3]. An analogous relation for $\bar{h}(k,\lambda)$, not explicitly stated here, can be derived from [5, Lemma 3b].

Lemma 3: A labeling λ of order $m\geq 2$ is an expanded cyclic Gray code if and only if $\bar{d}(1,\lambda)=\bar{d}(1,\beta_m)$ and $\bar{d}(3,\lambda)=\bar{d}(3,\beta_m)$.

Proof: First, that a labeling λ is a cyclic Gray code if and only if $\bar{d}(1,\lambda)=\bar{d}(1,\beta_m)$ follows from the definition of a cyclic Gray code. Second, that a cyclic Gray code λ of order $m\geq 3$ is an expanded cyclic Gray code if and only if $\bar{d}(3,\lambda)=\bar{d}(3,\beta_m)$ was proved in [5, Lemma 3]. The case $m=2$ is trivial. □

The critical index set depends on the modulation form and the order m, but the method used to find the critical index set for PSK and PAM is the same. We derive the critical index set for PSK in detail and only point out the essential differences in the derivation of the critical index set for PAM. QAM is not treated in this section, as it will be shown in Section VII that the concept of critical indices is not needed to analyze the performance of two-dimensional BRGC's.

A. The Critical Index Set for PSK

Theorem 4: For PSK constellations, the critical index sets of orders $m\leq 4$ satisfy $\Psi_1=\varnothing$, $\Psi_2=\Psi_3=\{1\}$, and $\Psi_4=\{1,3\}$.

Proof: It is trivial that $\Psi_1=\varnothing$. For $m\geq 2$, we consider the set $\Lambda_m(1)$ of labelings *not* having the cyclic Gray property. For all $m\geq 2$, there is at least one labeling in this set, so definitely $1\in\Psi_m$.

TABLE II
THE NUMBER OF BINARY GRAY CODES AND BINARY CYCLIC GRAY CODES THAT DO NOT HAVE IDENTICAL ADS AS A FUNCTION OF THE ORDER m. THE TABLE, WHICH WAS OBTAINED BY COMPUTER SEARCH, DOES NOT COUNT THE SAME ENTITIES AS [5, TAB. I], ALTHOUGH THE NUMERICAL VALUES AGREE FOR $m \leq 4$.

m	cyclic Gray	Gray
1	1	1
2	1	1
3	1	3
4	9	131

TABLE III
THE CRITICAL INDEX SET Ψ_m FOR PSK AND PAM AS A FUNCTION OF m.

m	Ψ_m (PSK)	Ψ_m (PAM)
1	$\varnothing$	$\varnothing$
2	$\{1\}$	$\{1\}$
3	$\{1\}$	$\{1,3\}$
4	$\{1,3\}$	$\{1,3,5\}$
5	$\{1,3,5\}$	$\{1,3,5,9\}$
6	$\{1,3,5,9\}$	$\{1,3,5,9,17\}$
7	$\{1,3,5,9,17\}$	$\{1,3,5,9,17,33\}$
8	$\{1,3,5,9,17,33\}$	$\{1,3,5,9,17,33,65\}$

The set $\Lambda_m(k)$ contains for $k \geq 2$ only cyclic Gray codes. However, *all* cyclic Gray codes have $\bar{d}(2) = 2$. Therefore, all cyclic Gray codes have identical ADS's for $k = 1$ and 2 and, hence, $\Lambda_m(2) = \varnothing$ for all orders m.

For $\Lambda_m(3)$, we turn to column 2 of Table II, where the number of cyclic Gray codes that do not have identical ADS is listed. Since there is only one cyclic Gray code for $m \leq 3$, $\Lambda_m(3) = \varnothing$ for $m \leq 3$, which completes the proof that $\Psi_2 = \Psi_3 = \{1\}$.

For $m = 4$, we conclude from Lemma 3 and the fact that the Gray code of order $m = 3$ is unique that there is only one expanded Gray code of order 4, which is the BRGC. Hence, $\Lambda_4(i) = \varnothing$ for $i \geq 4$. On the other hand, there exist several Gray codes of order 4 that are not the BRGC, see Table II, and consequently not expanded, which proves that $\Lambda_4(3) \neq \varnothing$. $\square$

Theorem 5: For $m \geq 5$, Ψ_m is obtained from Ψ_{m-1} by adding another element to the set, namely,

$$\Psi_m = \Psi_{m-1} \cup \{2(\max \Psi_{m-1}) - 1\}. \tag{30}$$

Proof: From the previous proof, $1 \in \Psi_m$ and $2 \notin \Psi_m$ for $m \geq 2$. To establish when $3 \in \Psi_m$, we recall from Lemma 3 that $\Lambda_m(3)$ is the set of Gray codes that are not expanded Gray codes. We will prove that this set is nonempty for $m \geq 4$ by showing that it includes the class of *balanced* Gray codes. Such labelings exist for all orders, see [12] and [13, pp. 14-15][1] and they have the property that the M bit transitions in a cyclic list of the M labels are distributed as evenly as possible among the m bit positions. To be precise, no more than $M/m + 2$ transitions occur in any one position of a balanced Gray code. In an expanded Gray code, on the other hand, half of the transitions occur in the same bit [5] (cf. Table I). Since $M/2 > M/m + 2$ for $m \geq 4$, we conclude that for $m \geq 4$, balanced Gray codes are not expanded, $\Lambda_m(3) \neq \varnothing$, and $3 \in \Psi_m$.

Now, for $T(\lambda) \geq 4$, we are dealing with the class of cyclic Gray codes for which $\bar{d}(k)$ is identical to the ADS of β_m for $k = 1, 2$, and 3. From Lemma 3, we know that all such labelings of order m can be constructed by expansion of a Gray code of order $m-1$. Hence, their ADS's can be calculated using (28)–(29). From the recursions we find that the critical index $T(\lambda_{m-1})$ of a labeling of order $m-1$ will propagate to the expanded labeling λ_m and result in a critical index

$$T(\lambda_m) = 2T(\lambda_{m-1}) - 1.$$

In summary, $\Psi_m = \{1,3\} \cup \{2i-1 \mid i \in \Psi_{m-1}\}$ for $m \geq 5$, which is equivalent to (30). $\square$

The PSK critical index sets Ψ_m are listed in the second column of Table III for $m = 1, \ldots, 8$.

[1] The perhaps earliest proof of the existence of balanced Gray codes for all m is attributed to T. Bakos [14].

B. The Critical Index Set for PAM

For PAM, the critical index set is derived in a similar way; the difference is that the ADS is defined by (18) and we exclude the cyclic requirement on the Gray codes. The change of definition for the ADS results in different values for $\bar{h}(k)$ for $k = 1, 2$ compared to the PSK case, but the conclusion that $\bar{h}(k)$ is identical for all Gray codes of order $m \geq 3$ for $k = 1, 2$ is still valid [5, Lemma 2b].

From column 3 of Table II, we see that in the class of not necessarily cyclic Gray codes, there are three classes of Gray codes of order $m = 3$ that do not have identical ADS's. This means that the critical index sets are not the same for PAM as for PSK. Indeed, they are given by the following two theorems.

Theorem 6: $\Psi_1 = \varnothing$, $\Psi_2 = \{1\}$, and $\Psi_3 = \{1, 3\}$.

Theorem 7: For $m \geq 4$, Ψ_m is given by (30).

We omit the proofs, which are analogous to the proofs of Theorems 4–5. The critical index sets of PAM are listed in the third column of Table III for $m = 1, \ldots, 8$.

V. THE OPTIMAL PSK LABELING

At this point, we have established the foundation required to address the proof of the optimality of β_m. In this section, we use the results in Section IV to derive sufficient conditions on the signal energy-to-noise ratio for which β_m is optimal for an M-PSK system.

A. The Bounding Ratio

The procedure we use is to compare β_m to all labelings in Λ_m. This is done by finding, for each $i \in \Psi_m$, a signal energy-to-noise ratio γ such that

$$P_{\mathrm{b}}(\beta_m, \gamma) \leq \min_{\lambda \in \Lambda_m(i)} P_{\mathrm{b}}(\lambda, \gamma) \tag{31}$$

and showing that the highest of these values taken over all $i \in \Psi_m$ provides a γ above which β_m yields the lowest possible BEP over the Gaussian channel.

For this purpose, we define the *bounding ratio* for PSK as

$$R(i,\gamma) \triangleq \frac{\sum_{k=i+1}^{M-i-1} P(k,\gamma)}{2P(i,\gamma)} \tag{32}$$

whose significance is given by the following lemma.

Lemma 8: For any $i \in \Psi_m$, a sufficient criterion for (31) is that

$$R(i,\gamma) \leq \frac{2}{M(m-1)}.$$

Proof: Let $\epsilon \triangleq 2/M$ and $\omega \triangleq m-1$. Define

$$\check{d}_i(k) \triangleq \begin{cases} \bar{d}(k,\beta_m), & k = 0,\ldots,i-1 \\ \bar{d}(k,\beta_m)+\epsilon, & k = i \\ \bar{d}(k,\beta_m)-\omega, & k = i+1,\ldots,M/2 \\ \check{d}_i(M-k) & k = M/2+1,\ldots,M-1. \end{cases} \tag{33}$$

The value at $k = i$ is a lower bound on the difference between the ADS of a labeling $\lambda \in \Lambda_m(i)$ and β_m. To show this, we observe that the sum (8) for any given k contains the same number of terms for which $\boldsymbol{c}_l$ has odd parity and $\boldsymbol{c}_{(l+k) \bmod M}$ has even parity as vice versa. In both cases, the Hamming distance is odd, whereas in all other cases, it is even. Hence, (8) contains an even number of odd terms, which proves that the resolution of $\bar{d}(k,\lambda)$ is ϵ. From this fact and (2), we conclude that the ADS of any $\lambda \in \Lambda_m(i)$ satisfies $\bar{d}(k,\lambda) \geq \check{d}_i(k)$.

If now $R(i,\gamma) \leq \epsilon/\omega$, then

$$\begin{aligned} 0 &\leq 2P(i,\gamma)\epsilon - \sum_{k=i+1}^{M-i-1} P(k,\gamma)\omega \\ &= \sum_{k=1}^{M-1} P(k,\gamma)\left(\check{d}_i(k) - \bar{d}(k,\beta_m)\right) \\ &\leq \sum_{k=1}^{M-1} P(k,\gamma)\bar{d}(k,\lambda) - \sum_{k=1}^{M-1} P(k,\gamma)\bar{d}(k,\beta_m) \\ &= P_{\mathrm{b}}(\lambda,\gamma) - P_{\mathrm{b}}(\beta_m,\gamma) \end{aligned}$$

for any labeling $\lambda \in \Lambda_m(i)$. □

Note that the bounds given by ϵ and ω in (33) are chosen for their simplicity; it is possible to find and use tighter bounds, but we have yet to find bounds that would give more than a marginal effect on the derived upper thresholds.

B. BRGC Optimality Thresholds for M-PSK

We now proceed to derive sufficient conditions of optimality of β_m for M-PSK over the Gaussian channel. We will evaluate (32) for each $i \in \Psi_m$ and find a range of γ for which all these $|\Psi_m|$ inequalities are valid simultaneously.

Lemma 9: For any $m \geq 3$ and $i \in \Psi_m$ and with a_i and b_i as defined in (9), there exists a unique $\gamma = \gamma_m(i)$ that satisfies the inequality

$$\frac{Q\left(\sqrt{2\gamma}\sin a_i\right)}{Q\left(\sqrt{2\gamma}\sin b_i\right)} \geq 1 + \frac{M(m-1)}{2} \tag{34}$$

with equality. The inequality is valid for all $\gamma \geq \gamma_m(i)$.

Proof: The left-hand side of (34) is equal to one for $\gamma = 0$ and a continuous function of γ for $\gamma \geq 0$. To complete the proof, we will show that it is also strictly increasing and unbounded. From the implicit definition in (9),

$$\max_{i\in\Psi_m} b_i = b_{\max\Psi_m} = \frac{(2\max\Psi_m + 1)\pi}{2^m}.$$

Since $\max\Psi_m = 2^{m-3}+1$ for $m \geq 4$ and $\max\Psi_3 = 1$,

$$b_i \leq \left(\frac{1}{4} + \frac{3}{2^m}\right)\pi < \frac{\pi}{2}$$

for any $m \geq 3$ and $i \in \Psi_m$. From Lemma 2, we see that for $0 \leq a_i < b_i \leq \pi/2$, i.e., $0 \leq \sin a_i < \sin b_i$, the left-hand side of (34) is strictly increasing with $\sqrt{2\gamma}$ (and therefore also with γ). In addition, invoking well-known bounds on the Q-function [6, p. 98], we have

$$\frac{Q\,(ax)}{Q\,(bx)} \geq \frac{b}{a}\left(1 - \frac{1}{a^2x^2}\right)e^{x^2(b^2-a^2)/2} \tag{35}$$

which, for $b > a > 0$, can be made arbitrarily large by increasing x. □

The value $\gamma_m(i)$ defined in this lemma is the threshold above which the BRGC of order m is better than any labeling in $\Lambda_m(i)$, as stated in the following theorem.

Theorem 10: $P_{\mathrm{b}}(\lambda,\gamma) \geq P_{\mathrm{b}}(\beta_m,\gamma)$ for every $m \geq 2$, $i \in \Psi_m$, $\gamma \geq \gamma_m(i)$, and $\lambda \in \Lambda_m(i)$.

Proof: For $m = 2$, to begin with,

$$R(1,\gamma) = \frac{P(2,\gamma)}{2P(1,\gamma)} < \frac{1}{2}$$

so Lemma 8 is satisfied for all γ and $i \in \Psi_2 = \{1\}$. Hence, β_2 is optimal for order $m = 2$ at any SNR.

For $m \geq 3$, the bounding ratio (32) is rewritten using (9)–(10) as

$$R(i,\gamma) = \frac{\Gamma\,(b_i,\gamma)}{\Gamma\,(a_i,\gamma) - \Gamma\,(b_i,\gamma)} \tag{36}$$

which is valid for $1 \leq i \leq M/2-1$. In general, the bounding ratio is tedious to handle directly, so we derive a more tractable upper bound on $R(i,\gamma)$ using the Q-function. Again referring to Figure 1, an upper bound on $\Gamma(a,\gamma)$ for $0 \leq a \leq \pi/2$ is

$$\Gamma\,(a,\gamma) \leq 2Q\left(\sqrt{2\gamma}\sin a\right).$$

Furthermore, for $0 \leq a \leq b \leq \pi/2$, the difference $\Gamma(a,\gamma) - \Gamma(b,\gamma)$ is lowerbounded by

$$2Q\left(\sqrt{2\gamma}\sin a\right) - 2Q\left(\sqrt{2\gamma}\sin b\right) \leq \Gamma\,(a,\gamma) - \Gamma\,(b,\gamma).$$

Now (36) yields, for all $i \geq 1$ such that $0 \leq a_i \leq b_i \leq \pi/2$,

$$\begin{aligned} R(i,\gamma) &\leq \frac{Q\left(\sqrt{2\gamma}\sin b_i\right)}{Q\left(\sqrt{2\gamma}\sin a_i\right) - Q\left(\sqrt{2\gamma}\sin b_i\right)} \\ &= \frac{1}{Q(\sqrt{2\gamma}\sin a_i)/Q(\sqrt{2\gamma}\sin b_i) - 1}. \end{aligned}$$

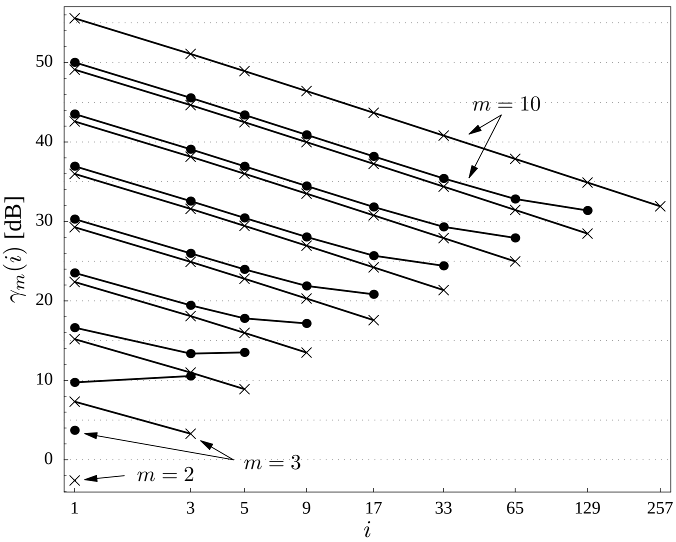

Fig. 2. The function $\gamma_m(i)$ for $i \in \Psi_m$ from $m = 2$ (bottom, PAM only) to $m = 10$ (top pair of curves). Bullets (•) mark M-PSK and crosses (×) mark M-PAM.

For $\gamma \geq \gamma_m(i)$, the denominator is at least $M(m-1)/2$ by Lemma 9. Lemma 8 completes the proof. □

Corollary 11: For any M-PSK constellation, the optimal labeling at asymptotically high SNR is the BRGC.

In Figure 2, the function $\gamma_m(i)$ is shown for $i \in \Psi_m$ and $m = 3, 4, \ldots, 10$. The interpretation of $\gamma_m(i)$ is the following. Consider a labeling $\lambda \in \Lambda_m(i)$, i.e., a labeling for which $T(\lambda) = i$. For $\gamma > \gamma_m(i)$, β_m will result in a lower BEP according to (7) over a Gaussian channel, irrespective of the ADS values of λ for $k > i$. If $\gamma < \gamma_m(i)$, there may exist labelings $\lambda \in \Lambda_m(i)$ such that the BEP is lower than for β_m even though $\bar{d}(i, \beta_m) < \bar{d}(i, \lambda)$. For example, we see from Figure 2 that for $m = 4$, any cyclic Gray code will give lower BEP than any non-Gray labeling for $\gamma \geq 9.7$ dB. Compared to the cyclic Gray codes in $\Lambda_4(3)$, β_4 gives a lower BEP than all these labelings for $\gamma \geq 10.5$ dB. For 9.7 dB $\leq \gamma \leq 10.5$ dB, the optimal labeling may be different from β_4, but it must be a cyclic Gray code.

We define the *maximal optimality threshold* of order $m \geq 3$ as

$$\hat{\gamma}_m = \max_{i \in \Psi_m} \gamma_m(i) \tag{37}$$

and we let, formally, $\hat{\gamma}_2 = -\infty$. Clearly, $\hat{\gamma}_m$ is an upper bound on the corresponding optimality threshold γ_m^*. The main theorem for PSK now follows from Theorem 10.

Theorem 12: The BRGC is optimal for any M-PSK constellation at $\gamma \geq \hat{\gamma}_m$.

The maximal optimality threshold (as seen in Figure 2) is given in Table IV for $m = 2, 3, \ldots, 10$. We note that for this range of m, it is only for $m = 4$ that $\hat{\gamma}_m \neq \gamma_m(1)$. From the column $P_\mathrm{b}(\hat{\gamma}, \beta_m)$, computed as detailed in Section VIII, we conclude that the maximal optimality thresholds are indeed quite low from a practical viewpoint; the BRGC is the optimal labeling whenever the targeted BEP P_b is less than 1.6 % and $m \leq 10$, which covers most M-PSK systems of practical interest.

TABLE IV

THE M-PSK MAXIMAL OPTIMALITY THRESHOLDS $\hat{\gamma}$ FOR $m = 2, \ldots, 10$ FROM (37) AND LEMMA 9, THE CORRESPONDING BIT ENERGY-TO-NOISE RATIO $\hat{\gamma}_\mathrm{b} \triangleq \hat{\gamma}/m$, AND THE BEP WHEN THE BRGC IS USED AT $\gamma = \hat{\gamma}$.

m	$\hat{\gamma}$ [dB]	$\hat{\gamma}_\mathrm{b}$ [dB]	$P_\mathrm{b}(\beta_m, \hat{\gamma})$
2	$-\infty$	$-\infty$	0.5
3	3.8	-1.1	0.147
4	10.6	4.6	0.090
5	16.7	9.7	0.070
6	23.6	15.8	0.049
7	30.3	21.9	0.036
8	37.0	28.0	0.027
9	43.6	34.0	0.021
10	50.1	40.1	0.016

VI. THE OPTIMAL PAM LABELING

In this section, we apply the methods used to prove optimality of β_m for M-PSK to systems using M-PAM.

A. The Bounding Ratio

The proof method for the PAM case is very similar to that of the PSK case; the main difference lies in the evaluation of the crossover probabilities. We again use the inequality (31), this time using (15) for the BEP expression, in order to find a signal energy-to-noise ratio threshold above which β_m gives the lowest possible BEP of all labelings.

Lemma 13: For any $i \in \Psi_m$,

$$\frac{\sum_{k=i+1}^{\infty} \mathcal{P}(k, \mu)}{\mathcal{P}(i, \mu)} \leq \frac{1}{(m-1)(M-1)} \tag{38}$$

is a sufficient criterion for $P_\mathrm{b}(\beta_m, \gamma) \leq \min_{\lambda \in \Lambda_m(i)} P_\mathrm{b}(\lambda, \gamma)$.

Proof: We rewrite (18) for $k \geq 1$ as

$$\bar{h}(k) = \frac{1}{2M} \left(\sum_{j=0}^{M-2} d_\mathrm{H}(\boldsymbol{c}'_j, \boldsymbol{c}'_{j+k}) + \sum_{j=1}^{M-1} d_\mathrm{H}(\boldsymbol{c}'_j, \boldsymbol{c}'_{j-k}) \right) \tag{39}$$

and observe that each of the $2M-2$ terms is between 1 and m, inclusively. Hence,

$$1 - \frac{1}{M} \leq \bar{h}(k) \leq 1 - \frac{1}{M} + \omega$$

where $\omega \triangleq (m-1)(1-1/M)$. Furthermore, the resolution of $\bar{h}(i)$ is $\epsilon \triangleq 1/M$, which can be shown by considering the following two cases separately. If $i = 1$, the terms of $\bar{h}(i)$ in (39) are pairwise equal, and if $i > 1$, λ is a Gray code and the sum for $\bar{h}(i)$ contains an even number of odd values. In both cases, $2M\bar{h}(i)$ is even.

Therefore, the ADS of any labeling $\lambda \in \Lambda_m(i)$ satisfies $\bar{h}(k, \lambda) \geq \check{h}_i(k)$, where

$$\check{h}_i(k) \triangleq \begin{cases} \bar{h}(k, \beta_m), & k = 0, 1, \ldots, i-1 \\ \bar{h}(k, \beta_m) + \epsilon, & k = i \\ \bar{h}(k, \beta_m) - \omega, & k = i+1, i+2, \ldots \end{cases}$$

If now (38) holds, then

$$\begin{aligned} 0 &\le \mathcal{P}(i,\mu)\epsilon - \sum_{k=i+1}^{\infty} \mathcal{P}(k,\mu)\omega \\ &= \sum_{k=1}^{\infty} \mathcal{P}(k,\mu)\left(\check{h}_i(k) - \bar{h}(k,\beta_m)\right) \\ &\le \sum_{k=1}^{\infty} \mathcal{P}(k,\mu)\bar{h}(k,\lambda) - \sum_{k=1}^{\infty} \mathcal{P}(k,\mu)\bar{h}(k,\beta_m) \\ &= P_{\mathrm{b}}(\lambda,\gamma) - P_{\mathrm{b}}(\beta_m,\gamma) \end{aligned}$$

for any labeling $\lambda \in \Lambda_m(i)$. □

As for the PSK case, it is possible to sharpen these bounds in many ways, e.g., by letting $\check{h}_i(k) = \bar{h}(k,\beta_m)$ for $k \ge M-1$, but as such improvements appear to influence the overall BEP very little, we use the above bounds for simplicity.

B. BRGC Optimality Thresholds for M-PAM

The theorems in this Section are analogous to similar theorems in Section V-B, but the proofs are simpler, thanks to the attractive properties of the Gaussian Q-function.

Lemma 14: For any $m \ge 2$ and $i \in \Psi_m$, there exists a unique $\gamma = \gamma_m(i)$ that satisfies

$$\frac{Q\left((2i-1)\mu(\gamma)\right)}{Q\left((2i+1)\mu(\gamma)\right)} \ge 1 + (m-1)(M-1) \tag{40}$$

for $i \ge 1$ with equality. The inequality is valid for all $\gamma \ge \gamma_m(i)$.

Proof: The left-hand side of (40) is equal to one for $\gamma = 0$ and a continuous function of γ for $\gamma \ge 0$. From Lemma 2 and the relation (16) between μ and γ, the left-hand side of (40) is strictly increasing in γ for a given i. It can be made arbitrarily large, as shown in (35). □

Theorem 15: $P_{\mathrm{b}}(\lambda,\gamma) \ge P_{\mathrm{b}}(\beta_m,\gamma)$ for every $m \ge 2$, $i \in \Psi_m$, $\gamma \ge \gamma_m(i)$, and $\lambda \in \Lambda_m(i)$.

Proof: The theorem follows immediately from Lemma 13, (17), and Lemma 14. □

Corollary 16: For any M-PAM constellation, the optimal labeling at asymptotically high SNR is the BRGC.

The solutions to (40) for $i \in \Psi_m$ are shown in Figure 2 for $m = 2, 3, \ldots, 10$. The maximal optimality thresholds, again defined as in (37), are listed in Table V along with the resulting $P_{\mathrm{b}}(\beta_m, \hat{\gamma}_m)$, computed as in Section VIII. In analogy with Theorem 12, the main result of this section is stated as a theorem, which follows immediately from Theorem 15.

Theorem 17: The BRGC is optimal for any M-PAM constellation at $\gamma \ge \hat{\gamma}_m$.

The last column of Table V indicates that the theorem holds for most M-PAM systems of practical interest ($m \le 10$ and $P_{\mathrm{b}} \le 1.5$ %). For $m = 2$, we also compare the upper bound $\hat{\gamma}_2$ with the optimality threshold γ_2^*, which can be calculated exactly. To do this, we first generate all distinct labelings (in the sense of having different ADS's) of order $m = 2$. There are three such labelings: the BRGC, the natural binary code (NBC), and another non-Gray labeling. We calculate their differential ADS's (22) and equate pairwise their BEP's (21) to find all intersections between the BEP curves. The result is $\gamma_2^* = 5\mu^2/2$

TABLE V

THE M-PAM MAXIMAL OPTIMALITY THRESHOLDS $\hat{\gamma}$ OBTAINED FROM (37) AND LEMMA 14, THE CORRESPONDING BIT ENERGY-TO-NOISE RATIO $\hat{\gamma}_{\mathrm{b}} \triangleq \hat{\gamma}/m$, AND THE BEP WHEN THE BRGC IS USED AT $\gamma = \hat{\gamma}$.

m	$\hat{\gamma}$ [dB]	$\hat{\gamma}_{\mathrm{b}}$ [dB]	$P_{\mathrm{b}}(\beta_m, \hat{\gamma})$
2	–2.6	–5.6	0.277
3	7.3	2.5	0.146
4	15.2	9.2	0.090
5	22.4	15.4	0.061
6	29.3	21.5	0.044
7	36.0	27.5	0.032
8	42.6	33.6	0.025
9	49.1	39.6	0.019
10	55.6	45.6	0.015

where μ is the positive root of $Q(\mu) - 3Q(3\mu) + 2Q(5\mu)$. For any $\gamma < \gamma_2^*$, the NBC is the best labeling, while for $\gamma > \gamma_2^*$, of course, the BRGC is the best one. At the threshold $\gamma_2^* = -5.22$ dB, the BEP is $P_{\mathrm{b}}(\beta_2, \gamma_2^*) = 0.337$, to be compared with $P_{\mathrm{b}}(\beta_2, \hat{\gamma}_2) = 0.277$.

VII. THE OPTIMAL QAM LABELING

Not surprisingly, a similar technique as in the previous two sections can be applied to rectangular QAM constellations. We will show that the same results hold: the two-dimensional BRGC is optimal for high enough SNR, and finite thresholds are obtained above which the BRGC is better than any other labeling. However, the QAM case is different from PSK and PAM in two respects. Firstly, the critical indices are irrelevant; if we only determine when Gray codes are better than non-Gray codes for QAM, then earlier results can be used to establish that the BRGC is the best of all Gray codes. Secondly, the maximal optimality threshold turns out to be much higher than for PSK and PAM.

The starting point is the BEP expression (24) and the ADS components $\bar{t}(k,l,\lambda)$ and $\bar{t}_0(k,\lambda)$. We will upperbound and lowerbound these components for Gray and non-Gray labelings, respectively. This allows us to lowerbound the difference in BEP for the two classes of labelings. Particular attention is paid to $\bar{t}_0(1,\lambda)$, which is the value where Gray codes differ from non-Gray labelings. It can be shown that for any rectangular Gray code λ_{G} of order (m_1, m_2),

$$\begin{aligned} \bar{t}_0(1,\lambda_{\mathrm{G}}) &= 1 - \frac{1}{2M_1} - \frac{1}{2M_2} \triangleq \hat{t}_1 \\ \bar{t}_0(k,\lambda_{\mathrm{G}}) &\le m_1 + m_2 \triangleq \hat{t}_2, & k \ge 2 \\ \bar{t}(k,l,\lambda_{\mathrm{G}}) &\le m_1 + m_2 \triangleq \hat{t}_3, & k,l \ge 1 \end{aligned}$$

and for any non-Gray labeling λ_{NG} with $m_1 \ge 2$ and $m_2 \ge 2$,

$$\begin{aligned} \bar{t}_0(1,\lambda_{\mathrm{NG}}) &\ge \hat{t}_1 + \epsilon \\ \bar{t}_0(k,\lambda_{\mathrm{NG}}) &\ge 1 - \frac{1}{2M_1} - \frac{1}{2M_2} = \hat{t}_2 - \omega_2, & k \ge 2 \\ \bar{t}(k,l,\lambda_{\mathrm{NG}}) &\ge 1 - \frac{1}{M_1 M_2} = \hat{t}_3 - \omega_3, & k,l \ge 1 \end{aligned}$$

TABLE VI

THE $M \times M$ QAM MAXIMAL OPTIMALITY THRESHOLDS $\hat{\gamma}$ FROM THEOREM 18, THE CORRESPONDING BIT ENERGY-TO-NOISE RATIO $\hat{\gamma}_{\mathrm{b}} = \hat{\gamma}/(2m)$, AND THE BEP WHEN THE BRGC IS USED AT $\gamma = \hat{\gamma}$.

m	$\hat{\gamma}$ [dB]	$\hat{\gamma}_{\mathrm{b}}$ [dB]	$P_{\mathrm{b}}(\beta_m \times \beta_m, \hat{\gamma})$
2	12.6	6.5	0.022
3	21.5	13.7	$2.7 \cdot 10^{-3}$
4	29.2	20.2	$3.9 \cdot 10^{-4}$
5	36.4	26.4	$6.3 \cdot 10^{-5}$
6	43.4	32.6	$1.1 \cdot 10^{-5}$
7	50.1	38.7	$2.0 \cdot 10^{-6}$
8	56.8	44.7	$3.8 \cdot 10^{-7}$
9	63.3	50.8	$7.5 \cdot 10^{-8}$
10	69.8	56.8	$1.5 \cdot 10^{-8}$

$\epsilon \triangleq 3/(2M_1M_2)$, $\omega_2 \triangleq m_1 + m_2 - 1 + 1/(2M_1) + 1/(2M_2)$, and $\omega_3 \triangleq m_1 + m_2 - 1 + 1/(M_1M_2)$ are all positive.

Theorem 18: Let

$$t(\mu) \triangleq (2\epsilon + \omega_3)Q(\mu) + (\epsilon + \omega_2)\frac{Q(3\mu)}{Q(\mu)}. \tag{41}$$

If $t(\mu(\gamma_G)) = \epsilon$, where the relation between μ and γ for QAM is given by (23), then $P_{\mathrm{b}}(\lambda_G, \gamma) \le P_{\mathrm{b}}(\lambda_{\mathrm{NG}}, \gamma)$ for any Gray code λ_G, any non-Gray labeling λ_{NG}, and any $\gamma \ge \gamma_G$.

Proof: From (24), the difference in BEP between any non-Gray labeling and any Gray code can now be lowerbounded as

$$\begin{aligned} P_{\mathrm{b}}(\lambda_{\mathrm{NG}}, \gamma) - P_{\mathrm{b}}(\lambda_{\mathrm{G}}, \gamma) &\ge \frac{4}{m_1 + m_2} \cdot \Bigg(\epsilon\mathcal{P}(0,\mu)\mathcal{P}(1,\mu) \\ &\quad - \omega_2\mathcal{P}(0,\mu)\sum_{k=2}^{\infty}\mathcal{P}(k,\mu) - \omega_3\sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\mathcal{P}(k,\mu)\mathcal{P}(l,\mu)\Bigg) \\ &= \frac{4}{m_1+m_2}\Big(\epsilon(1-2Q(\mu))(Q(\mu)-Q(3\mu)) \\ &\quad - \omega_2(1-2Q(\mu))Q(3\mu) - \omega_3 Q^2(\mu)\Big) \\ &\ge \frac{4}{m_1+m_2}(\epsilon - t(\mu))Q(\mu). \end{aligned}$$

By Lemma 2 and the monotonicity of $Q(\mu)$, $t(\mu(\gamma))$ is a continuous, monotonically decreasing function for $\gamma \ge 0$. It ranges from $t(0) = 2\epsilon + \omega_2 + \omega_3/2$ to $\lim_{\gamma\to\infty} t(\mu(\gamma)) = 0$ (see (35)). Thus, there is a unique positive value γ_G for which $t(\mu(\gamma_G)) = \epsilon$. For any $\gamma \ge \gamma_G$, $t(\mu(\gamma)) \le \epsilon$ and thus $P_{\mathrm{b}}(\lambda_{\mathrm{G}}, \gamma) \le P_{\mathrm{b}}(\lambda_{\mathrm{NG}}, \gamma)$. □

So far we have shown that all Gray codes are better than all non-Gray codes in rectangular QAM systems with $\gamma \ge \gamma_G$, but which Gray code is the best of them? This question can be answered without further analysis by exploiting known results.

Theorem 19: Let

$$\hat{\gamma} \triangleq \max\left\{\gamma_G, \left(1 + \frac{M_2^2 - 1}{M_1^2 - 1}\right)\hat{\gamma}_{m_1}, \left(1 + \frac{M_1^2 - 1}{M_2^2 - 1}\right)\hat{\gamma}_{m_2}\right\} \tag{42}$$

where $\hat{\gamma}_m$ are the maximal optimality thresholds for PAM. Then the BRGC is optimal for any $M_1 \times M_2$ QAM constellation at $\gamma \ge \hat{\gamma}$.

Proof: From [2], we know that the only way to assign a Gray code to a rectangular QAM constellation is by taking the direct product of two PAM constellations, each labeled by a Gray code. The BEP of such direct product constellations is given by (27), which is minimized for any $\gamma \ge \hat{\gamma}$ when $\lambda_1 = \beta_{m_1}$ and $\lambda_2 = \beta_{m_2}$, according to Theorem 17 and the definition (42). We conclude that the two-dimensional BRGC is optimal for QAM whenever $\gamma \ge \hat{\gamma}$. □

The maximal optimality thresholds $\hat{\gamma}$ of square constellations are listed in Table VI. They were obtained by numerically solving (41). The corresponding BEP of the two-dimensional BRGC, evaluated as in Section VIII, is also listed. It is interesting to observe that the upper thresholds are much higher than the corresponding values for PSK and PAM, and $\hat{\gamma} = \gamma_G$ for all orders m in the range of the tables. It is still safe to conclude that the BRGC is asymptotically optimal even for QAM, but we cannot claim that the BRGC is optimal in the SNR range of practical interest. This is partly due to the fact that the used bounding technique appears to be weaker for QAM than for PSK and PAM, but also to the fact that non-Gray codes are indeed more competitive in two dimensions. Specifically, the two most likely symbol errors for QAM require that the norm of the noise vector exceeds μ and $\mu\sqrt{2}$, respectively, whereas the corresponding values for PAM are μ and 3μ. Therefore, sacrificing the Gray property, which implies that more bit errors are associated with the most likely error pattern, carries a heavier penalty for PAM than QAM.

If the optimality threshold γ^*_{PAM} is known for an M-PAM constellation, then $\check{\gamma} = 2\gamma^*_{\mathrm{PAM}}$ is a lower bound on the optimality threshold for an $M \times M$ QAM constellation via (27). Specifically, we conclude from the 4-PAM results in Sec. VI-B that the two-dimensional BRGC is the best *product* labeling (but not necessarily the best labeling) for 4×4 QAM at $\gamma > \check{\gamma} = -2.21$ dB, for which $P_{\mathrm{b}}(\beta_2 \times \beta_2, \check{\gamma}) = 0.337$, and that the product of two NBC's is the best product labeling[2] at $\gamma < \check{\gamma}$. This lower bound, however, is still far from the upper bound for $m = 2$ in Table VI and we do not know where γ^* lies in this interval.

VIII. BEP OF SYSTEMS WITH BRGC LABELINGS

Now that the BRGC has been shown to minimize the BEP of multilevel PSK, PAM, and QAM transmission over the Gaussian channel for large enough SNR, we evaluate this minimum BEP in the three cases. This is achieved by deriving closed-form expressions for the differential ADS of the BRGC and utilizing the general BEP expressions given in Section III.

In [7], it was shown that if an M-PSK constellation is labeled by β_m, the ADS is given by

$$\bar{d}(k, \beta_m) = \mathrm{tri}\,(2^m, k) + \sum_{i=2}^{m}\mathrm{tri}\left(2^i, k\right) \tag{43}$$

for all integers k. The function $\mathrm{tri}(N, k)$ is a periodic triangular function of period N, defined by

$$\mathrm{tri}\,(N, k) \triangleq 2\left|\frac{k}{N} - \left\lfloor\frac{k}{N}\right\rceil\right|$$

[2] A third labeling that attains the same BEP at $\gamma = \check{\gamma}$ is the product of one BRGC and one NBC.

where the function $\lfloor x \rceil$ rounds x to the closest integer (ties are rounded arbitrarily)

To calculate the BEP in the form (13), we need the differential ADS of the BRGC. Since $\bar{d}(k)$ is a sum of triangular sequences, $\bar{\Delta}(k)$ is a sum of piecewise constant functions. In particular,

$$\operatorname{tri}(N,k) - \operatorname{tri}(N,k-1) = \frac{2}{N}(-1)^{\lfloor 2(k-1)/N \rfloor} \tag{44}$$

where $\lfloor x \rfloor$ denotes the integer part of x. Combining (14) with (43) and (44), we obtain

$$\bar{\Delta}(k,\beta_m) = \frac{(-1)^{\left\lfloor (k-1)2^{1-m} \right\rfloor}}{2^{m-1}} + \sum_{i=1}^{m-1} \frac{(-1)^{\left\lfloor (k-1)2^{-i} \right\rfloor}}{2^i} \tag{45}$$

for all integers k. We believe that (45), combined with (13), is the simplest published form for the exact BEP of M-PSK with the BRGC.

The BEP of PAM constellations with the BRGC can be computed using either (15) or (21). Since the former method turns out to yield somewhat complicated expressions [15, Pt. E], we treat in this paper only the latter method, which is based on the differential ADS. Thus, the BEP is given by (21) in combination with the following theorem, which is proved in the Appendix. Another expression for the BEP of PAM constellations was given in [11, eq. (9)–(10)], with a more complicated proof.

Theorem 20: The differential ADS of a PAM constellation labeled with the BRGC of order m is, for $1 \le k \le 2^m - 1$,

$$\bar{\delta}(k,\beta_m) = \sum_{i=1}^{m} \left(\frac{1}{2^i} - \frac{1}{2^m} \left\lfloor \frac{k-1/2}{2^i} \right\rceil \right) (-1)^{\left\lfloor (k-1/2)/2^i \right\rfloor}. \tag{46}$$

The two-dimensional BRGC is the direct product of two one-dimensional BRGC's. Thus, the BEP of a rectangular QAM system with the BRGC is simply given by (27) in combination with (21) and (46). A recursive method to compute the same BEP was given in [16].

IX. Conclusions and Comments

We have addressed the problem of finding an optimum signal constellation labeling with respect to minimizing the BEP for M-PSK, M-PAM, and $M_1 \times M_2$ QAM under the assumptions of a Gaussian channel, equally likely and statistically independent transmitted bits, and coherent maximum likelihood symbol detection. The result is that for the asymptotic case when the signal energy-to-noise ratio γ approaches infinity, the BRGC gives the lowest possible BEP among all Gray codes (and other labelings), for all three modulation types.

The BRGC is in fact the optimal labeling for a significant range of values for γ. In particular, the BRGC is shown to be optimal as long as $\gamma \ge \hat{\gamma}$, where $\hat{\gamma}$ is an upper bound on the optimality threshold γ^* (defined as the smallest SNR for which the BRGC yields the smallest possible BEP). Numerical values of $\hat{\gamma}$ are given, and by evaluating the BEP at the thresholds, the conclusion is drawn that when the BEP is below a few percent, the BRGC is the optimum labeling for PSK and PAM. The same conclusion cannot be drawn for QAM, possibly because the derived upper bounds on γ^* are too loose.

The paper includes new closed-form expressions for the BEP of the three modulation formats with *BRGC labelings* (Section VIII). These expressions, which we believe are the simplest available for the purpose, have the additional benefit of separating the influence of the labeling on the BEP from that of the constellation geometry. Analogous BEP expressions for *arbitrary labelings* are also given (Section III).

Appendix
Proof of Theorem 20

For labelings that are symmetric in the sense that $d_H(\boldsymbol{c}_0, \boldsymbol{c}_l) = d_H(\boldsymbol{c}_{M-1}, \boldsymbol{c}_{M-1-l})$ for all $l = 0, \ldots, M-1$, (22) can be simplified to

$$\bar{\delta}(k,\lambda) = \frac{1}{M} \sum_{l=0}^{M-k-1} \left(d_H(\boldsymbol{c}_l, \boldsymbol{c}_{l+k}) - d_H(\boldsymbol{c}_{l+1}, \boldsymbol{c}_{l+k}) \right). \tag{47}$$

Define a mapping $f : \{0,1\} \to \mathbb{Z}$ such that $f(0) = 1$ and $f(1) = -1$. If the components of a label $\boldsymbol{c}_i \in \{0,1\}^m$ are denoted $c_{i,m}, c_{i,m-1}, \ldots, c_{i,1}$, then

$$d_H(\boldsymbol{c}_j, \boldsymbol{c}_l) = \frac{1}{2} \sum_{i=1}^{m} \left(1 - f(c_{j,i}) f(c_{l,i})\right).$$

With this notation, (47) can be written as

$$\bar{\delta}(k,\lambda) = \frac{1}{2M} \sum_{i=1}^{m} \sum_{l=0}^{M-k-1} f(c_{l+k,i}) \left(f(c_{l+1,i}) - f(c_{l,i}) \right). \tag{48}$$

From Table I or by induction on m, it is easily verified that the labels $(\boldsymbol{c}_0, \ldots, \boldsymbol{c}_{M-1})$ of the BRGC satisfy

$$f(c_{l,i}) = (-1)^{\left\lfloor (l+1/2)2^{-i} \right\rceil}$$

for $l = 0, \ldots, M-1$ and $i = 1, \ldots, m$. Furthermore, since the BRGC possesses the required symmetry, (47) and (48) hold and the differential ADS of the BRGC is

$$\bar{\delta}(k,\beta_m) = \frac{1}{2M} \sum_{i=1}^{m} \sum_{l=0}^{M-k-1} (-1)^{\left\lfloor (l+k+1/2)2^{-i} \right\rceil} \cdot \left[(-1)^{\left\lfloor (l+3/2)2^{-i} \right\rceil} - (-1)^{\left\lfloor (l+1/2)2^{-i} \right\rceil} \right].$$

The bracketed expression is nonzero only when $l = (n + 1/2)2^i - 1$ for some integer n. When l goes from 0 to $M-k-1$, then n goes from 0 to $\hat{n} \triangleq \left\lfloor (M-k)/2^i - 1/2 \right\rfloor$. (If the bracketed expression is zero for all l in the interval, then $\hat{n} = -1$ and the sum over n below should be interpreted as zero.) Thus,

$$\bar{\delta}(k,\beta_m) = \frac{1}{2M}\sum_{i=1}^{m}\sum_{n=0}^{\hat{n}}(-1)^{n+\left\lfloor 1/2+(k-1/2)2^{-i}\right\rceil} \cdot \left[(-1)^{n+\left\lfloor(1+2^{-i})/2\right\rceil} - (-1)^{n+\left\lfloor(1-2^{-i})/2\right\rceil}\right]$$
$$= \frac{1}{2M}\sum_{i=1}^{m}(\hat{n}+1)(-1)^{\left\lfloor 1/2+(k-1/2)2^{-i}\right\rceil} \cdot \left[(-1)^{\left\lfloor(1+2^{-i})/2\right\rceil} - (-1)^{\left\lfloor(1-2^{-i})/2\right\rceil}\right]$$
$$= \frac{1}{M}\sum_{i=1}^{m}(\hat{n}+1)(-1)^{\left\lfloor(k-1/2)2^{-i}\right\rfloor}. \tag{49}$$

For any $0<\epsilon<1$ and any integers x and y, $\lfloor x/y+1/2\rfloor = \lfloor (x+\epsilon)/y\rceil$. Applying this identity to $\hat{n}$ and letting $\epsilon=1/2$, we obtain

$$\hat{n} = \left\lfloor\frac{M-k+1/2}{2^i}\right\rceil - 1$$

which substituted into (49) yields (46).

ACKNOWLEDGMENT

We thank two anonymous reviewers for their suggestions, which improved the presentation of the paper significantly.

REFERENCES

[1] J. G. Proakis, *Digital Communications*, McGraw-Hill, 3rd edition, 1995.
[2] R. D. Wesel, X. Liu, J. M. Cioffi, and C. Komninakis, "Constellation labeling for linear encoders," *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2417–2431, Sept. 2001.
[3] F. Gray, "Pulse code communications," U. S. Patent No. 2632058, Mar. 1953.
[4] G. Takahara, F. Alajaji, N. C. Beaulieu, and H. Kuai, "Constellation mappings for two-dimensional signaling of nonuniform sources," *IEEE Transactions on Communications*, vol. 51, no. 3, pp. 400–408, March 2003.
[5] 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, no. 12, pp. 3170–3182, Dec. 2004.
[6] S. Verdú, *Multiuser detection*, Cambridge University Press, 1998.
[7] J. Lassing, E. G. Ström, T. Ottosson, and E. Agrell, "Computation of the exact bit error rate of coherent M-ary PSK with Gray code bit mapping," *IEEE Transactions on Communications*, vol. 51, no. 11, pp. 1758–1760, Nov. 2003.
[8] C. E. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell System Technical Journal*, vol. 38, pp. 611–656, May 1959.
[9] J. Lassing, E. Ström, T. Ottosson, and E. Agrell, "The exact symbol and bit error probabilities of coherent M-ary PSK," in *Proc. IEEE International Symposium on Information Theory*, Yokohama, Japan, p. 11, June 2003.
[10] M. K. Simon and M. S. Alouini, *Digital Communications over Fading Channels. A Unified Approach to Performance Analysis*, John Wiley & Sons, 2000.
[11] K. Cho and D. Yoon, "On the general BER expression of one- and two-dimensional amplitude modulations," *IEEE Transactions on Communications*, vol. 50, no. 7, pp. 1074–1080, July 2002.
[12] G. S. Bhat and C. D. Savage, "Balanced Gray codes," *The Electronic Journal of Combinatorics*, vol. 3, no. 1, 1996. paper no. R25.
[13] D. E. Knuth, *The Art of Computer Programming*, vol. 4, fascicle 2. Addision-Wesley, 2005.
[14] C. Savage, "Comments by the second author on R25 of volume 3 (1), 1996," *The Electronic Journal of Combinatorics*, 2005.
[15] J. Lassing, "On the labeling of signal constellations," Ph.D. dissertation, Dept. of Signals and Systems, Chalmers University of Technology, Göteborg, Sweden, Jan. 2005.
[16] L.-L. Yang and L. Hanzo, "A recursive algorithm for the error probability evaluation of M-QAM," *IEEE Communications Letters*, vol. 4, no. 10, pp. 304–306, Oct. 2000.

Erik Agrell received the M.S. degree in electrical engineering in 1989 and the Ph.D. degree in information theory in 1997, both from Chalmers University of Technology, Sweden.

From 1988 to 1990, he was with Volvo Technical Development as a Systems Analyst, and from 1990 to 1997, with the Department of Information Theory, Chalmers University of Technology, as a Research Assistant. In 1997–1999, he was a Postdoctoral Researcher with the University of Illinois at Urbana-Champaign and the University of California, San Diego. In 1999, he joined the faculty of Chalmers University of Technology as an Associate Professor, holding a position first at Chalmers Lindholmen University College and since 2001 in the Department of Signals and Systems. His current research interests include geometrical aspects of coding theory, in particular lattices and block codes, bit-to-symbol mappings in digital communication systems, and coding and modulation for fiber-optic channels.

Dr. Agrell served as Publications Editor for IEEE TRANSACTIONS ON INFORMATION THEORY from 1999 to 2002.

Johan Lassing was born in Borås, Sweden, in 1973. He received the M.S. degree in engineering physics in 1995 and the Ph.D. degree in communication systems in 2005, both from Chalmers University of Technology, Sweden.

From 1997 to 1999 he was with the Swedish space system developer Omnisys Instruments AB and was part of the development of the scientific measurements satellite ODIN until its launch in 2001. In 2001 he started Qamcom Technology AB, a company devoted to communication systems equipment implementation, currently with 14 employees. Dr. Lassing currently works as the CEO of Qamcom Technology AB.

His main research interests lie in the fields of modulation and coding theory.

Erik G. Ström (S'93–M'95–SM'01) received the M.S. degree from the Royal Institute of Technology (KTH), Stockholm, Sweden, in 1990, and the Ph.D. degree from the University of Florida, Gainesville, in 1994, both in electrical engineering. He accepted a postdoctoral position at the Department of Signals, Sensors, and Systems at KTH in 1995. In February 1996, he was appointed Assistant Professor at KTH, and in June 1996 he joined the Department of Signals and Systems at Chalmers University of Technology, Göteborg, Sweden, where he is now a Professor in Communication Systems since June 2003 and head of the Communication Systems Group since 2005. He received the Chalmers Pedagogical Prize in 1998. Since 1990, he has acted as a consultant for the Educational Group for Individual Development, Stockholm, Sweden. He is a contributing author and associate editor for Roy Admiralty Publishers' FesGas-series, and was a co-guest editor for the special issue of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS on Signal Synchronization in Digital Transmission Systems, 2001. Dr. Ström was a member of the board of the IEEE VT/COM Swedish Chapter 2000-2006. His research interests include communication theory in general, and constellation labelings, channel estimation, synchronization, multiple access, and wireless sensor networks in particular. He has published more than 70 conference and journal papers.

Tony Ottosson was born in Uddevalla, Sweden, in 1969. He received the M.Sc. in Electrical Engineering from Chalmers University of Technology, Göteborg, Sweden, in 1993, and the Lic. Eng. and Ph.D. degrees from the Department of Information Theory, Chalmers University of Technology, in 1995 and 1997, respectively.

Currently he is a Professor at the Communication Systems Group, Department of Signals and Systems, Chalmers University of Technology. During 1999 he was also working as a Research Consultant at Ericsson Inc, Research Triangle Park, NC, USA. From 1995 to 1998 he participated in the European FRAMES (Future Radio wideband Multiple Access System) project both as a co-worker and activity leader of the area of coding and modulation. Dr. Ottosson has served as Associate Editor for IEEE Vehicular Technology from 2000 to 2004.

His research interests are in communication systems and information theory. Specific topics are modulation, coding, CDMA, multiuser detection, combined source-channel coding, joint decoding techniques, synchronization, cross-layer interaction and scheduling. Dr. Ottosson has published more than 90 journal and conference papers, and holds several patents.