A new concept of Bit Geometrical Uniformity and its application to BICM-ID systems

Trung Kien HOANG, Binh Minh NGUYEN, and The Cuong DINH
Dept. of Radio-Electronics Engineering, Le Qui Don Technical University, 100 Hoang Quoc Viet Avenue, Hanoi, Vietnam (e-mail: cuongdt@lqdtu.edu.vn)

Abstract— **A new concept of Bit Geometrical Uniformity (BGU) is introduced to help the analysis and estimation of performance of systems with Iterative Detection/Decoding (ID). It is also shown that the BGU labelings of the signal constellation allows a fast search for best component binary Recursive Systematic Convolutional (RSC) codes of a BICM-ID system using both of overall and in-line interleaving. Best rate-1/2 RSC codes are reported for 4PSK constellation as an example. The new concept and new search method can be directly extended to MPSK and MQAM constellations.**

Index Terms—**Geometrical Uniformity, Bit Geometrical Uniformity, Bit-Interleaved Coded Modulation.**

I. Introduction

GEOMETRICAL uniformity together with Geometrically Uniform (GU) signal constellation, GU binary labelings and related GU codes were first introduced by Forney [1] as an intelligent generalization and good extension of earlier works of Calderbank and Sloane [2], Benedetto [3]. The linearity of the coding systems and the regularity of the signal constellation in used were combined to provide the so called uniform error property (UEP). With UEP analysis and design of signal space codes are simplified, because only the all-zero sequence must be considered to evaluate any kind of code performance.

Geometrical uniformity then was extended to design multidimentional trellis coded modulation (TCM) codes (see [4], [5] and references therein). However, there were at least two problems to be solved. Firstly, it is complex to apply the concept of GU in analysis and design of an important class of coded modulation systems, namely, the widely interested bit-interleaved coded modulation (BICM) schemes without or with iterative decoding (ID) due to the interleaving at bit level and the general mapping to the signal constellation. Secondly, for most transmission systems involving the binary information sequences that must be transmitted to the receiver, the bit error probability (BEP) is more important performance measure than the symbol sequence error probability (SEP). To solve the second problem, Garello *et al.* [6] has introduced the concept of bit geometrically uniform (BGU) labelings which should help to design encoders with the so called uniform bit error property (UBEP). Though [6] reported some important theoretical results on BGU, the definition of the BGU labelings was so strict that there were only few BGU labelings with usual constellations (PSK, PAM, and QAM). The application of BGU was serially concatenated trellis codes with UBEP [6].

The definition of BGU labelings given in [6] was related to maximum likelihood (ML) decoding. To release this very strong condition in order to widen the use of the UBEP concept, we note that the complex ML decoding/detection has been recently replaced by sub-optimum iterative decoding/detection. At high enough SNRs performance of ID comes very close to the performance of ML decoding due to the ideal feedback (IF) [8]. Furthermore, in research of BICM-ID systems most of efforts is concentrated on the design of signal constellations and signals' labelings (see [7] and references therein) to obtain the best BER over Gaussian as well as Rayleigh channels. The interleaver is supposed to be random and, hence, must have a length of thousands bits (normally more than 5000). More interestingly, the topic is how to select the RSC code for a certain signal constellation with a certain mapping rule. Since bits in the label of a signal point are protected differently, the going on argument is about which output bit, the systematic bit or parity bits, of the RSC code must be protected better.

In this paper we introduce a new concept of BGU based on the use of ID, assuming that the feedback is perfect. New BGU labelings allow simplified analysis and estimation of performance of BICM-ID systems. Theoretical results of the paper are given without proof. Interested readers can contact authors for the full version of the paper. These results which appeared in the form of some upper bound on the error probabilities of BICM-ID systems are used in a full search for good component binary RSC codes, to end up the argument on which bit we have to protect better.

II. BGU labelings

Consider a finite set $S = \{s_0, s_1, \ldots, s_{M-1}\}$ of M signal points ($M = 2^m$). Suppose all signal points are used equally likely for the transmission over an additive white Gaussian noise (AWGN) channel. The average SER with ML decoding is

$$P_e(S) = \frac{1}{M}\sum_{i=0}^{M-1} P_e(s_i)\ ,$$

where the symbol error probability $P_e(s_i)$, when $s_T = s_i$ is transmitted and s_R is received, is $P_e(s_i) = P[s_R \neq s_T / s_T = s_i]$.

Definition 1: A set (constellation) S is said to satisfy the UEP if SEP with ML symbol decoding does not depend on the transmitted signal, i.e., $P_e(s_i)$ is the same for each $s_i \in S$.

Given a *signal set* S is a set of discrete points in an N-dimensional Euclidean space $\mathbf{R}^N$. An *isometry* u of $\mathbf{R}^N$ is a mapping of $\mathbf{R}^N$ onto itself that preserves the *Euclidean distance* as $\| u(\mathbf{x}) - u(\mathbf{y}) \|^2 = \| \mathbf{x} - \mathbf{y} \|^2$, where $u(\mathbf{x})$ denotes the image of $\mathbf{x}$ under the transformation u. All isometries of $\mathbf{R}^N$ can be derived from the three "primitive" transformations: *translation* (in a certain direction), *rotation* (about a certain line or axis), and *reflection* (relative to a certain hyperplane). For a given isometry u of $\mathbf{R}^N$ and a set $S \in \mathbf{R}^N$, we denote $u(S) = \{u(\mathbf{x}) : \mathbf{x} \in S\}$ the image of S under u. A *symmetry* of S is an isometry u that leaves S invariant, that is $u(S) = S$. All symmetries of S form a group $\Gamma(S)$, with respect to composition operation, called the *symmetry group* of S. Then a signal set S is *geometrically uniform* (GU) if, given any two points $\mathbf{x}$ and $\mathbf{y}$ in S, there exists a symmetry u of S such that $u(\mathbf{x}) = \mathbf{y}$. We also say that S is GU if it is generated from a point $\mathbf{x} \in S$ under the action of its symmetry group. The subgroup of $\Gamma(S)$ that is minimally sufficient to generate S from its arbitrary point is called the *generating group* of S.

It is well known that if S is GU then the UEP holds [1]. Most of usual signal sets are either GU (MPSK, permutation alphabets, generalized permutation alphabets [8]) or approximately GU if we neglect the boundary effect (QAM) .

Given a finite constellation S with cardinality $M = 2^m$, a binary labeling $\mu[S,m]$ for S is a one-to-one function μ that associates a distinct m-bit label $c = \mu(s)$ to each signal $s \in S$.

To evaluate the bit error probability, suppose that the binary information sequences are partitioned into frames of length m, and that all frames have the same probability. When a frame c_T must be transmitted, the signal $s_T = \mu^{-1}(c_T)$ is sent over the AWGN channel. After ML symbol decoding, one obtains the received signal s_R and the received frame $c_R = \mu(s_R)$. The bit error probability with ML symbol decoding, when a signal s_i (or a frame c_i) is transmitted, is then

$$P_b(s_T = s_i) = \sum_{j \neq i} P_b(s_R = s_j / s_T = s_i)$$
$$= \sum_{j \neq i} \frac{w_H(\mu(s_j) \oplus \mu(s_i))}{m} P(s_R = s_j / s_T = s_i)$$

The average BEP over with ML symbol decoding is given by

$$P_b(S) = \frac{1}{M} \sum_{i=0}^{M-1} P_b(s_T = s_i)$$

Definition 2: A binary labeling $\mu[S,m]$ for a constellation S is said to satisfy the UBEP if the bit error probability with ML symbol decoding does not depend on the transmitted signal, i.e., $P_b(s_T = s_i)$ is the same for each signal $s_i \in S$.

Given a GU constellation S with generating group G, there is a one-to-one correspondence between an element $g_i \in G$ and a signal $s_i = g_i(s_0) \in S$. For simplicity, we will denote the labeling function $\mu(g_i(s_0))$ as $\mu(g_i)$, and we will assume $\mu(e) = 0$, where e is the group identity and 0 is the m-bit all-zero label.

Theorem 1: Let $\mu[S,m]$ be a binary labeling of a GU constellation S with generating group G. Then μ has the UBEP if it satisfies the following distance rule:

$$d_H(\mu(g_i), \mu(g_j)) = w_H(\mu(g_i^{-1} \bullet g_j)) \quad \forall g_i, g_j \in G \qquad (1)$$

A labeling satisfying Theorem 1 is called a *bit geometrically uniform* (BGU) labeling [6]. Note that the existence of a BGU labeling is a sufficient (not necessary) condition for UBEP, like geometrical uniformity is only a sufficient condition for UEP.

III. A NEW CONCEPT OF BGU LABELINGS

We consider again a finite constellation S with cardinality $M = 2^m$ and a binary labeling $\mu[S,m]$. For a particular signal point $s_i \in S$ let $c_i = (c_i^1, c_i^2, \ldots, c_i^m)$, $i = 0, \ldots, M-1$ be the binary label such that $c_i = \mu(s_i)$. Let s_i and s_{i*} be a pair such that $c_i^j \oplus c_{i*}^j = 1$, $c_{i*} = \mu(s_{i*})$. We define the j-th bit distance associated with a signal $s_i \in S$ as the Eulidean distance $d_E(s_i, j) = \|s_i - s_{i*}\|$. Obviously, $d_E(s_i, j) = d_E(s_{i*}, j)$. In general, $d_E(s_i, j)$ depends on the signal s_i and there may be up to $M/2$ different $d_E(s_i, j)$ for each j.

Definition 3: A binary labeling $\mu[S,m]$ for a constellation S is a *wide-sense BGU labeling* if for each $j \in \{1, \ldots, m\}$ the bit distance does not depend on the signal, i.e., $d_E(s_i, j)$ is the same for each signal s_i, say, $d_E(s_i, j) = d_E^j$.

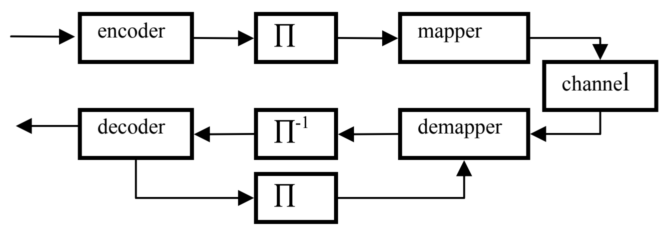

Fig. 1. System model

In [8], this property the labeling is called IF-regularity. We consider a BICM-ID system as depicted in Fig. 1 with a feedback from the channel encoder to the demapper for iterative demapping and decoding. The bit interleaving has two types: the overall interleaving and the in-line interleaving (see Fig.2). Although the in-line interleavers form only a subset of overall interleavers, they have some nice properties that can be employed to improve the system performance.

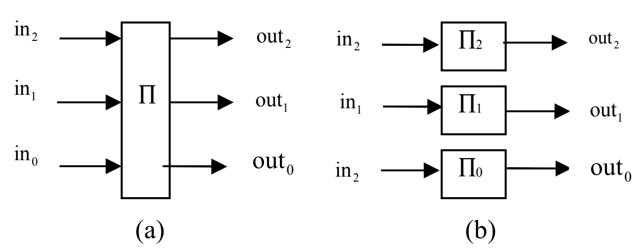

Figure 2. Overall (a) and in-line (b) bit interleavers.

A sequence of K information bits is encoded by a binary RSC code whose $N = K/R$ output bits are bit-interleaved by a random interleaver π. Here R is the rate of the encoder. At time index t, the bits $c_t = (c_t^1, c_t^2, \ldots, c_t^m), t = 1, \ldots, N/m$, from the interleaved sequence are mapped to a symbol $s_t \in S$ by the labeling $c_t = \mu(s_t)$. The interleaving length and the number of iterations are set large enough, namely $N=10^4$ and $I=20$, so that it is possible to assume an IF. In this case, the AWGN channel with M-ary modulation, $M=2^m$, can be seen as m separate parallel binary channels [8]. If the labeling is BGU, each of these binary channels is characterized by a bit distance. Thus, the j-th channel, $j\in\{1,\dots,m\}$, has a bit distance d_E^j and an equivalent bit energy $E_b^j=(d_E^j)^2/4$.

Example: In this paper we let $R=1/2$, $S=4PSK$. Consider the set partition (SP) labeling of the 4PSK constellation in use (Fig. 3). It can be seen that this labeling satisfies the condition of Definition 3 and, hence, is BGU. We distinguish two BGU labelings. The labeling SP-A has $d_E^1=2\sqrt{E_S}$, $d_E^2=\sqrt{2E_S}$ and the labeling SP-B has $d_E^1=\sqrt{2E_S}$, $d_E^2=2\sqrt{E_S}$, where E_S is the symbol energy of the 4PSK signal.

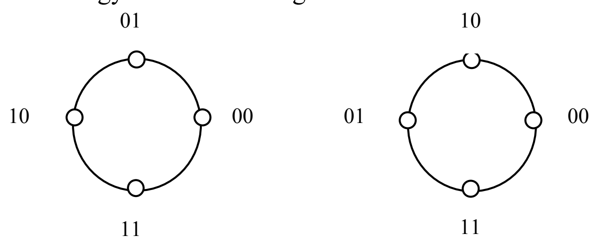

Fig. 3. BGU labelings of the 4PSK constellation

In the overall bit interleaving system, under the assumption of the BGU labeling and IF, each bit in the interleaver's output sequence selects (with probability $1/m$) for transmission a binary channel of energy $E_b^j=(d_E^j)^2/4$, $j\in\{1,\dots,m\}$. The average modulation energy then is $[(d_E^1)^2+\cdots+(d_E^m)^2]/4m$. Transmission of a symbol now can be then imagined as the consecutive m uses of the average binary channel, each lasts $T_c=T_S/m$ seconds, where T_S is the symbol duration. The equivalent noise bandwidth is m times as large as the original transmission bandwidth. If N_0 is the noise density of the BICM-ID system in consideration,then the noise density of the equivalent binary channel is $N_0^*=N_0/m$.

Remark 1: Under the assumption of IF and BGU labelings the performance of the BICM-ID system in error floor range is equivalent to the ML performance of the component RSC over an equivalent AWGN channel with BPSK modulation of a signal-to-noise power ratio (SNR)

$$\frac{E_b}{N_0^*}=\frac{1}{4RN_0}\sum_{j=1}^{m}(d_E^j)^2$$

For a large information bit frame (a large K), without loss of generality we can assume a terminated to the all-zero state RSC code whose codewords in fact form a linear block code of rate K/N. Since interleaving is a linear transformation of each codeword, at the interleaver's output we have another block code. We can state the following remark.

Remark 2: The linearity of the code and the BGU labeling of the signal constellation (by Definition 3) implies that the UBEP holds for the BICM-ID in consideration.

IV. Code Construction

A. The BICM-ID system with overall interleaving

Theorem 2: Consider a BICM-ID system with overall bit interleaving and rate-k/n RSC code and transmission over an AWGN channel with one-side noise power spectral density N_0. Under the assumption of IF and BGU labelings, the BEP of the system is upper bounded as follows

$$P_b<\frac{1}{k}\frac{\partial T(D,I,L)}{\partial I}\bigg|_{I=1,L=1} \quad (1)$$

where the dummy variable D is given as

$$D=exp\left\{-\frac{1}{4N_0}\sum_{j=1}^{m}(d_E^j)^2\right\} \quad (2)$$

First, we note that the bound (1) belongs to a class of so called transfer function upper bounds, so one can apply tightening techniques and numerical computation as it has been described in [9]. *Second*, the bound (1) differs from the classical Viterbi bound only in computation of the dummy variable D, more exactly, in the exponent which represents the SNR. Thus, one can approximately compute the gain of the BICM-ID system compared with performance of the RSC code over an AWGN channel with BPSK and SNR equal to

$$\frac{E_S}{mR(N_0/m)}=\frac{E_S}{RN_0}=\frac{E_b}{N_0} \quad (3)$$

where mR is the number of information bits sent by a symbol of energy E_S. *Third*, good codes for AWGN channels will perform well in BICM-ID systems with overall interleaving.

Depicted curves in Fig.4 compare the upper bounds (1) with simulation results of the BICM-ID system for 4-state and 64-state RSC codes. Bounds have the same behaviour as curves of simulation results have, and expected to be good approximatin of BER at high SNRs.

B. Upper bound on BEP of BICM-ID system with in-line interleaving

With in-line interleaving the encoder output bits are not completely decoupled so that each component bit sequence is assigned to a certain binary channel among m parallel channels formed by the M-ary modulation, given IF condition holds.

In this paper we consider 4PSK modulation. Thus we have $M=4,\ m=2$. We use rate-1/2 RSC codes. Without loss of generality, we associate systematic bits c_t^s with the bit distance d_E^1 and parity bits c_t^p with the bit distance d_E^2. Consider an error event of length L associsted with binary sequence

$$C_L=c_1^sc_1^p,c_2^sc_2^p,\dots,c_L^sc_L^p$$

We have $W_H(C_L)=W_H(C_L^s)+W_H(C_L^p)$, where C_L^s and C_L^p are systematic and parity subsequences, respctively. When transmitted over equivalent binary channels, the error event sequence has a squared Euclidean distance to the all-zero path

$$(d_L)^2=W_H(C_L^s)(d_E^1)^2+W_H(C_L^p)(d_E^2)^2$$
$$=W_H(C_L)\frac{(d_E^1)^2+(d_E^2)^2}{2}+\frac{[W_H(C_L^s)-W_H(C_L^p)][(d_E^1)^2-(d_E^2)^2]}{2}$$

It can be seen that the first term is the average squared Euclidean distance of the same error event when the overall interleaving is used. Interestingly, if the code is matched to the labeling of the signal constellation so that the second term is positive, then there is a gain over the system with overall interleaving. With $m=2$ we have $N_0^* = N_0/2$.

Remark 3: Equivalent AWGN channels with BPSK modulation for transmission systematic bits and parity bits respectively have following SNRs

$$\frac{E_b^s}{N_0^*} = \frac{1}{2RN_0}(d_E^1)^2 \text{ and } \frac{E_b^p}{N_0^*} = \frac{1}{2RN_0}(d_E^2)^2$$

Theorem 3: Consider a BICM-ID system with in-line bit interleaving and rate-$1/2$ RSC code and transmission over an AWGN channel with one-side noise power spectral density N_0. Under the assumption of IF and BGU labelings, the BEP of the system is upper bounded as follows

$$P_b < \sum_{L=L_{\min}} \sum_{i=1}^{N_L} W_H(C_{L,i}^s) D^{W_H(C_{L,i}^s)(d_E^1)^2 + W_H(C_{L,i}^p)(d_E^2)^2} \quad (4)$$

where N_L is the number of length-L error events and $D = \exp\{-1/(2N_0)\}$.

The bound (4) is computed numerically by the technique presented in [9]. For each of memmory lengths $\upsilon = 2,3,\ldots,9$, the search is done for labelings SP-A and SP-B. The best code found for these two labelings are compared to obtain the best combination of the RSC code and the labeling in the sense of a lower value of the bound (4). For the 4PSK constellation it seems that the labeling SP-A always gives rise to better systems. RSC encoder generators in octal are given in Tab. 1. Bounds and simulation results are presented in Fig.4.

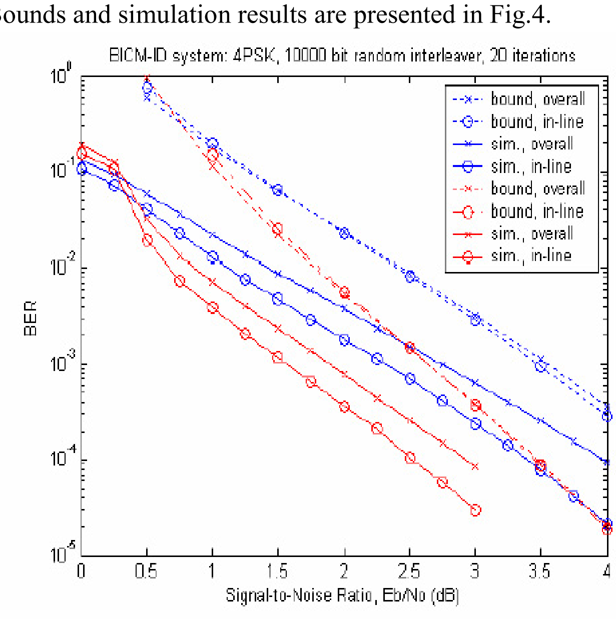

Fig. 4. Comparison of upper bounds and simulation results
Blue: RSC code [1,5/7]; Red: RSC code [1,35/23]

Table 1. Best RSC encoders found with the labeling SP-A

υ	$E_b/N_0 = 2$dB g_1 g_2	$E_b/N_0 = 4$dB g_1 g_2	$E_b/N_0 = 6$dB g_1 g_2
2	7 5	7 5	7 5
3	13 17	13 17	15 17
4	23 35	23 35	31 27
5	75 53	73 51	73 51
6	115 177	147 135	147 135
7	351 277	373 225	373 225
8	731 523	467 625	467 625
9	1073 1745	1073 1745	1561 1237

V. CONCLUSIONS

Based on the new concept of BGU labeling of GU signal constellation used in BICM-ID systems, new upper bounds on BEP of the BICM-ID system are reported together with code search results for in-line interleaving and 4PSK. It seems that the encoder and the labeling must combine to give the best performance of the BICM-ID system with in-line interleaving. The result can be extended to the cases of 8PSK and M-QAM constellations.

REFERENCES

[1] G. David Forney, “Geometrically uniform codes,” *IEEE Trans. Inform. Theory*, vol. 37, pp. 1241-1260, No. 5, Sept. 1991.

[2] A. R. Calderbank and N. J. A. Sloane, “New trellis codes based on lattices and cosets,” *IEEE Trans. Inform. Theory,* vol. IT-33, pp. 177-195, Jan. 1987.

[3] S. Benedetto, M. A. Marsan, G. Albertengo, and E. Giachin, “Combined coding and modulation: theory and applications,” *IEEE Trans. Inform. Theory* Vol. 34, pp. 223-236, March 1988.

[4] S. Benedetto, R. Garello, M. Mondin, and G. Montorsi, “Geometrically uniform partitions of $L \times MPSK$ constellations and related binary trellis codes,” *IEEE Trans. Inform. Theory*, vol. 39, pp. 1773-1798, No. 5, Nov. 1993.

[5] T. C. Dinh and T. Hashimoto, “A systematic approach to the construction of bandwidth-efficient multidimentional trellis codes,” *IEEE trans. Commun.*, vol. 48, no. 11, pp. 1808-1817, Nov. 2000.

[6] R. Garello, G. Montorsi, S. Benedetto, D. Divsala, and F. Pollara, “Labelings and encoders with the uniform bit error property with application to serially concatenated trellis codes,” *IEEE Trans. Inform. Theory*, vol. 48, No. 1, pp. 123-136, Jan. 2002.

[7] F. Schreckenbach, P. Henkel, N. Gortz, and G. Baud, “Analysis and design of mappings for iterative decoding of BICM,” in *Proceedings of XI National Symposium on Radio Sciences, Poznan, 7-8 April, 2005,* Poznan, Germany.

[8] Anders Nilsson, and Tor M. Aulin “On in-line bit interleaving for serially concatenated systems,” in Proceedings of IEEE International Conference on Communication (ICC), Seoul, South Korea, May 2005.

[9] Nguyen Binh Minh and Dinh The Cuong, “A Tight upper bound on the bit error probability of convolutional codes,” 8th Vietnam conference on Radio & Electronics (REV02), Nov. 2002.