Trellis Coded Modulation with Partially Overlapped Signal Subsets

Van Khan TRAN[1], The Cuong DINH[1], and Takeshi HASHIMOTO[2]
1. Dept. of Radio-Electronics Engineering, Le Qui Don Technical University, 100 Hoang Quoc Viet Avenue, Hanoi, Vietnam (e-mail: cuongdt@lqdtu.edu.vn)
2. Dept. of Electronics Engineering, University of Electro-Communication, 1-5-1 Chofugaoka, Chofu-shi, Tokyo, Japan (e-mail: hasimoto@itl.ee.uec.ac.jp)

Abstract— **We present a simple yet effective scheme of Trellis Coded Modulation (TCM) codes using partially overlapped signal subsets as a shaping technique. The optimum enlargement of the QAM signal constellation is reported. We show that 2-state TCM with 4 partially overlapped signal subsets has a coding gain of 2.5 dB in comparison with uncoded transmission, which is about 0.5 dB higher than the Ungerboeck TCM code with the same number of states. The construction is then extended to any number of trellis states using 4-, 8-, 16-way partitions of the QAM constellation.**

Index Terms— **TCM codes, shaping the constellation, overlapping**

I. INTRODUCTION

Trellis coded modulation (TCM) [1] has found wide applications in digital transmission due to its high utilization of the available bandwidth. To send n bits/sym conventional TCM schemes use a signal constellation of 2^{n+1} signal points. When using the quadrature amplitude modulation (QAM), the constellation expansion ratio (CER) equals to 2, meaning the loss of about 3 dB in signal energy [4]. This loss should be traded with the increase in the minimum squared distance (MSD) between signal sequences.

A problem in TCM design with QAM signals is the peak-to-average power ratio (PAR). The traditional approach to increase the total gain of the TCM scheme is to use a circular signal constellation, rather than a square one, and this also reduces the PAR. It is well known in the theory of digital transmission that capacity of an additive white Gaussian noise (AWGN) channel is attained for the input with the Gaussian distribution. In fact, the input signal constellation realized by shaping techniques, like shell mapping and trellis shaping. Maximally, a gain of 1.5 dB can be obtained by shaping [3].

Another approach is to use overlapped signal subsets, the idea first introduced by Soleymani *et. al.* [5]. They used an analytical representation of trellis codes [6] in which the output signal for a state transition is calculated by a non-linear function of the current and the past L source symbols. If the non-linear function yields the same signal for different state transitions, then it is said that an overlapping occurs. The constellation in actual use now is reduced, meaning that the CER becomes less than 2. With a large number of signals, it is difficult to construct/compute non-linear mapping function.

In this paper, we propose a TCM scheme with re-mapping of a signal point of high energy to a point of low energy. In Sec. 2, we propose as an example a simple scheme of 2-state TCM code using partially overlapped signal subsets which shows about 0.5 dB coding gain better than the conventional Ungerboeck-type code with the same number of states. The performance analysis for TCM with partially overlapped signal subsets is given in Sec.3, where we show the existence of the optimum constellation expansion factor, which is the best trade-off attained between saving of average signal energy, thanks to re-mapping, and degradation in the performance of the code, due to the partial overlap. In Sec.4, we generalize the construction of TCM with partially overlapped signal subsets to any number of trellis states, using 4- and 8-way geometrically uniform (GU) [2] partition of the QAM constellation. We tabulate the code search result and compare new codes with Ungerboeck-type codes in term of transfer function bound on the first event error probability.

II. CONSTRUCTION OF GU CODES FROM 4 PARTIALLLY OVERLAPPED SIGNAL SUBSETS

We consider transmission at n bits/sym over the AWGN channel using QAM signals. The QAM signal constellation we design consists of $(1+p)2^n$ signal points in the lattice Z^2 translated by (1/2, 1/2), where $(1+p)$ for $0<p\le 1$ is the CER and $p=1$ corresponds to the case of conventional Ungerboeck-type TCM codes.

A. Example of Ungerboeck's Code

Fig.1 shows an example the best 2-state TCM encoder given in [1], together with the trellis diagram of the code. This code can be used to send at 5 bits/sym. The MSD of the code equals to 3 and is 3 times as large as the MSD of 32-QAM constellation for uncoded transmission. However, the expansion from 32-QAM to 64-QAM means a loss of 3.22 dB in signal energy. The net coding gain is then 1.55 dB. The 4-way partition of the 64-QAM constellation is shown in Fig.3. We can refer to the subset A as the translate of the sub-lattice $2Z^2$. The subset C is obtained by reflecting A about the axis Oy. Subsets B and D are results of reflecting of A and C, respectively, about the origin. These subsets correspond to the lattice partition $Z^2/RZ^2/2Z^2$ with MSD chain $1/2/4$ [4], where the notation R is expressed with a matrix

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

For the Ungerboeck-type code, all the signal points are used.

B. Construction of partially overlapped subsets

Let p be such that $0<p<1$ and $p2^n$ is an integer. The original 2^{n+1}-point signal constellation S now is divided into three sub-constellations: the *inner subconstellation* consisting of $(1-p)2^n$ signal points of lowest energies; the *outer subconstellation* consisting of $(1-p)2^n$ signal points of highest energies; the subconstellation of *remaining* signal points. The idea here is that outer signal points are, if

978-1-4244-2681-2/08/$25.00 ©2008 IEEE

selected by Ungerboeck's scheme, re-mapped to inner signal points. Then the signal points actually sent over the channel form a new constellation S' of $(1+p)2^n$ points.

For the QAM constellation discussed above, remapping can be carried out in the following manner. According to the above subdivision, each of the subsets A, B, C and D of S is divided into *inner part, outer part*, and the *rest*, respectively. We remap the outer *A*-points to inner *B*-points, and outer *B*-points to inner *A*-points. The similar remapping is applied to subsets *C* and *D*. Then we have four subsets of S' as follows. The subset A' consists of *A*-points of S' and inner *B*-points. The subset B' consists of *B*-points of S' and inner *A*-points. In the same way we can define subsets C' and D'. Clearly, $A' \cap B' \neq \phi$ and $C' \cap D' \neq \phi$. This construction reserves symmetries of the constellation and hence gives rise to GU codes.

C. Coding with Overlapped Signal subsets and Decoding

Fig.2 shows an example of the TCM coding with overlapped signal subsets constructed with the above remapping rule. We choose $n = 5$ bits/sym and $p = 0.25$. The constellation S' consists of 40 signal points, with 24 inner points delimited by the solid line in Fig.3.

The TCM encoder using the overlapped signal subsets is constructed as follows. S can be seen as a union of 2^{n-1} shells each consisting of 4 equal-energy points *a, b, c,* and *d* belonging to A, B, C, and D, respectively. Each point of *S* is labeled by a binary tuple $\underline{Z} = (Z_{n+1}, Z_n, \cdots, Z_2, Z_1)$. Two least significant bits Z_2, Z_1 identify a subset between A, B, C, or D. Bits $(Z_{n+1}, Z_n, \cdots, Z_3)$ identify a shell. We use the following re-mapping rule. The outer most shell is remapped to the inner most shell, the outer shell of lower energy is remapped to the inner shell of higher energy, and so on. In each shell, an *A*-point is remapped to a *B*-point, a *B*-point is remapped to an *A*-point. The same remapping rule is applied to *C*- and *D*-points.

To ease to the remapping, we enumerate shells in the order of increasing energy using the shell number $m = 0, 1, \cdots, 2^{n-1}-1$. The 2^{n-2} shells of the basic 2^n-point constellation that could be used for uncoded transmission has $Z_{n+1} = 0$. Remaining shells (in the expanded part) has $Z_{n+1} = 1$. Bits $(Z_n, \cdots, Z_3)$ are selected so that

$$m = Z_{n+1}\left(2^{n-1}-1\right) + (-1)^{Z_{n+1}} \sum_{i=3}^{n} Z_i 2^{i-3} \qquad (1)$$

With this labeling, the remapping of an outer point to an inner point now is a simple inversion of bits Z_{n+1} and Z_2. The block named "re-mapper" in Fig.2 takes in bits $(Z_{n+1}, Z_n, \cdots, Z_3)$ from the conventional Ungerboeck-type encoder and computes the shell number *m*. For a given expansion ratio *p*, if $2^{n-2} \leq m < (2-p)2^{n-2}$, then it means that the output of Ungerboeck-type encoder selects one of $(1-p)2^{n-2}$ outer shell which should be remapped to a corresponding inner shell, simply by outing 1 which is added (modulo 2) to bits Z_2 and Z_{n+1} to obtain Z_2' and Z_{n+1}'.

Like the Ungerboeck-type TCM code, the decoding is carried out by the Viterbi algorithm (VA) and the underlying trellis of the encoder (Fig.3). For each symbol in the received sequence, the VA decoder computes MSDs from the symbol to subsets A', B', C', and D', stores the points in A', B', C', and D' that are in the MSDs to the received symbol, and traces the path that accumulates the least squared distance. The last survived path is used together with binary labels of points in A', B', C', and D', stored at each decoding stage, to give the final decision on the information bit sequence. The decoder inverses bits Z_2' and Z_{n+1}' when it is needed to produces correct information bits Z_2 and Z_{n+1}.

III. Performance analysis

It is well known that the performance of a TCM code can be evaluated by the first event error probability of the code, P_e. When the code is GU, we can assume without any loss of generality that a sequence of signal points corresponding to the all zero state sequence is sent. Then it is convenient to compare different codes by using their transfer function upper bounds on P_e,

$$P_e < T(D)\big|_{D=\exp(-E_s/4N_0)} \qquad (2)$$

where E_s is the average signal, N_0 is the one-side spectral density of the Gaussian noise, and $T(D)$ is the transfer function of the code which is computed from the error state diagram [2].

Fig.4 shows the error state diagram of the 2-state Ungerboeck's code. Let us denote by S_{ij} the signal subset assigned to the trellis transition from the state *i* to the state *j*, with $i, j = 0,1$. For a state pair $(i, j) \neq (0,0)$, the transition from a state *i* to a state *j* in the error state diagram is labeled by the branch metric $N(d_{ij})D^{d_{ij}^2}$, where d_{ij}^2 is the MSD between S_{ij} and S_{00}, and $N(d_{ij})$ denotes the average number of signal points in S_{ij} that are at d_{ij}^2 to a point in S_{00}. When $(i, j) = (0,0)$, we have d_{00}^2 as the MSD of S_{00}, and $N(d_{00})$ as the average number of signal points in S_{00} that are at the MSD to a point in S_{00}. Using the technique described in [2], we have, for the 2-state Ungerboeck's code,

$$T(D) = 4D^4 + \frac{8D^3}{1-2D} \qquad (3)$$

For the new 2-state code with overlapped signal subsets (see Fig.4 for the trellis diagram), the label of a transition depends on whether the signal point of S_{00}, selected randomly by the uncoded input bits, belongs to the overlapped part or not. From the construction of overlapped signal subsets described on Sec. 2 we know that $S_{00} = A'$ consists of 2^{n-1} different signal points, among which $(1-p)2^{n-1}$ inner points are overlapped with $S_{11} = B'$. If an inner point is selected (with probability $(1-p)$), then $d_{00}^2 = 2$, $N(d_{00}) = 4$, and $d_{11}^2 = 0$, $N(d_{11}) = 1$. If the selected point is not an inner point (with probability *p*), we have $d_{00}^2 = 4$, $N(d_{00}) = 4$ and $d_{11}^2 = 2$, $N(d_{11}) = 4$. Since $S_{00} = A'$ does not overlap with $S_{01} = D'$ and $S_{10} = C$, the selection of a point in $S_{00} = A'$ does not affect the interset distance, but it does change the number of points at that distance. We have $d_{01}^2 = d_{10}^2 = 1$ and $N(d_{01}) = N(d_{10}) = 4(1-p) + 2p = 2(2-p)$. The error state diagram of the 2-state code with overlapped signal subsets is shown in Fig.4. The corresponding transfer function is

$$T(D) = 4(1-p)D^2 + 4pD^4 + \frac{4(2-p)^2 D^2}{p(1-4D^2)} \qquad (4)$$

Obviously, when *p* is small the probability of choosing an inner point is large, meaning the distance property of the code is worsened. However, the small expansion factor gives rise to smaller average energy of the signals actually used in transmission.

Lemma 1: Let *S* be a QAM constellation with circular boundary that consists of *M* signal points with the average energy *E*. Then the constellation with overlapped signal subsets constructed on a circular QAM signal constellation with $(1+p)M$ signal points, $0 \leq p \leq 1$, has approximately (for large *M*) an average energy $(1+p^2)E$.

By scaling the average signal energy actually transmitted over the channel to the average signal energy in uncoded transmission we have $D = \exp\left\{-E_s/(4(1+p^2)N_0)\right\}$, with *p*=1 for the Ungerboeck type code, and $0 < p < 1$ for the code using overlapped subsets. Fig.5 shows the behavior of the transfer function upper bounds for both cases, as the function of p, given $E_s / N_0 = 14$ dB. It can be seen that at this value of the SNR the new code outperforms the Ungerboeck-type code for $0 < p < 0.5$. To keep the symmetries of the constructed overlapped constellation, we choose p = 0.25. Fig.5 also shows the comparison of transfer function upper bounds on P_e of the 2-state Ungerboeck's code and new code for p = 0.25. At high SNR the new coding scheme gives a larger coding gain in comparison with Ungerboeck's code, by about 0.55 dB.

IV. Generalization to the Construction of GU Codes with 8 Partially Overlapped Signal Subsets

It is well known that the performance of a TCM code is mainly determined by its free distance, especially at high SNRs. The free distance of the TCM code is, in turn, limited by the minimum intraset distance of the signal subset assigned to a trellis transition. In other words, the minimum distance between signals assigned to parallel transitions (forming the error event of length 1) upper bounds the free distance. Obviously, to attain a larger coding gain than the 2-state scheme described in Sec. II, we need a finer partition of the QAM constellation such that the MSD of the subset at the highest partition level becomes as large as possible.

A. GU Partially Overlapped Signal Subsets

Let S be a GU signal set with a generating group G and let s_0 be an arbitrary point of S. If we can find a normal subgroup N of G such that the quotient group G/N is isomorphic to $(\mathbf{Z}_2)^k$ for some integer k, then $S = G(s_0)$ and G/N induces a GU partition $G(s_0)/N(s_0)$ which allows a binary isometric labeling $m:(\mathbf{Z}_2)^k \to G(s_0)/N(s_0)$ for $\mathbf{z} = (z_k, \cdots, z_2, z_1) \in (\mathbf{Z}_2)^k$. Let $N(s_0)$ be the subset generated by the normal subgroup N from an arbitrary, but fixed, point of S. An ordered set $g = \{g_1, g_2, \ldots, g_k\}$ with $g_i \in G$ for $1 \le i \le k$ represents the binary partition chain of S if any coset X of $N(s_0)$ in S can be expressed as $X = g_1^{z_1} g_2^{z_2} \cdots g_k^{z_k}(N(s_0))$, where $g_i^{z_i} = e$, the identity map, if $z_i = 0$ and $g_i^{z_i} = g_i$ if $z_i = 1$. The vector $\mathbf{z} = (z_k, \cdots, z_2, z_1)$ is the binary isometric label of the coset X.

To construct 2^k GU partially overlapped signal subsets for $k \ge 2$, we consider again the QAM signal constellation S with 2^{n+1} signal points. Then we can divide S into 2^{n+1-k} shells so that each consists of 2^k signal points of roughly equal energy each belonging to one of subsets in the 2^k-way partition of S. Each point of S is labeled by a tuple $\underline{Z} = (Z_{n+1}, Z_n, \cdots, Z_2, Z_1)$, among which k less significant bits identify one of 2^k signal subsets and the remaining bits identify a shell. For a given p, we remap $(1-p)2^{n-k}$ outer shells to $(1-p)2^{n-k}$ inner shells as follows. We remap the outer most shell to the inner most shell, the outer shell of lower energy to the inner shell of higher energy, and so on. In each shell, the point of the subset X in the 2^k-way partition of S is remapped to a point of $g_k(X)$, where g_k is the isometry used at the partition level k.

For $k = 2$ and $n = 5$ we partition the 64-point QAM constellation S into 4 subsets A, B, C, and D by using $g_1 = v_y$, the reflection in axis Oy, and $g_2 = v_y v_x$, the reflection in the origin of the coordinate system (see Example of Ungerboeck code, Sec.2). For $p = 0.25$, there are 6 outer most shells which should be remapped to 6 inner most shells (the 24 signal points delimited by the solid line in Fig.4a).

For $k = 3$, the 8-way partition of the QAM constellation is induced by the group partition chain $G = V^2 Tr(2\mathbf{Z}^2)/\{e, g_2\} Tr(2\mathbf{Z}^2) / Tr(2\mathbf{Z}^2)/\{e, t_{(2,2)}\} Tr(4\mathbf{Z}^2) = N$, where $t_{(2,2)} = t_{(2,0)} t_{(0,2)}$ is defined as a translation of a point $s = (x, y)$ to a point $(x+2, y+2)$. Then we have $\{v_x, v_x v_y, t_{(2,0)}\}$ representing the 8-way partition chain with the distance chain 1/2/4/8. With the 8-way partition of the QAM constellation, it is impossible to allocate a shell of 8 points a, b, c, d, e, f, g, h of equal energy which, at the same time, represent 8 subsets *A, B, C, D, E, F, G, H* respectively, in the 8-way partition. The requirement of equality of signal energy is then relaxed in forming 8 shells each consisting of all representatives of 8 subsets. For $k = 3$ (Fig.4b), the dashed line shows the boundary of the 40-point constellation with 8 partially overlapped subsets. The resulted 40-point constellation used in TCM with 8 overlapped signal subsets is composed of 5 inner most shells and differs slightly from the 40-point constellation in the case of 4 subsets.

B. Code Search

Since the construction of partially overlapped subsets in the above reserves the geometrical uniformity, we can apply the fast code search algorithm by computing the transfer function upper bound on the first event error probability [4].

The results of code search are given in Tables 1 and2 for the Ungerboeck-type codes and for codes with partially overlapped sets, with 4- and 8-way partition, respectively. The generators of the encoder are given in octal. To compare found codes with the Ungerboeck-type codes we draw curves of the transfer function upper bounds on the first event error probability of codes in Fig.6 and 7 for 4- and 8-way partition of the QAM constellation, respectively.

First, we note on the Ungerboeck-type codes codes found in our search which include all codes reported previously [1]. This confirms that the search algorithm based on the minimization of the upper bound on the first event error probability can give best codes. For codes of Table 1 (Fig.6) the error probability decreases slowly with the increase of the number of trellis states. This is explained by the small MSD of the subsets in the 4-way partition that limits the free distance of these codes. When the number of subsets in the partition is larger (Table 2 and Fig.7), the error probability decreases with the increasing in the number of trellis states.

The same tendency of error probability curves is observed for codes with partially overlapped subsets. However, we should note that, for a small number of trellis states, new codes give a better performance than the Ungerboeck-type codes. At low SNRs, codes with partially overlapped subsets perform better than the Ungerboeck-type codes of the same number of states. When the number of trellis states is small, error events have short lengths and it is difficult for them to accumulate distances. At this point, thus, saving in signal energy is more effective than increasing signal distance, and this is the advantage of the codes using overlapped signal sets.

V. Conclusions

We have presented a new technique for the construction of TCM codes with small constellation expansion ratio, based on partially overlapped signal subsets drawn from an original QAM constellation. Constellation shaping by re-mapping high-energy signal points to low-energy signal points gives rise to some TCM codes that have better performance in comparison with conventional Ungerboeck-type TCM codes of the same coding and decoding complexity.

References

[1] Ungerboeck, "Channel coding with multilevel/phase signals", *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 55-67, Jan. 1982.

[2] G. D. Forney, Jr., "Geometrically uniform codes", *IEEE Trans. Inform. Theory*,

vol. 37, pp. 1241-1260, Sept. 1991.
[3] U. Washmann, R. F. H. Fisher, and J. B. Huber, "Multilevel codes: Theoretical concepts and practical design rules", *IEEE Trans. Inform. Theory*, vol. 45, pp. 1361-1391, July 1999.
[4] D. T. Cuong and T. Hashimoto, "A systematic approach to the construction of bandwidth-efficient multidimensional trellis codes," *IEEE Trans. Commun.*, vol. 48, no. 11, pp. 1808-1817, Nov. 2000.
[5] M, R, Soleymani and L. Kang, "TCM schemes with partially overlapped signal constellations," *IEEE Trans. Commun.*, March 1993.
[6] A. R. Calderbank and J. E. Mazo, "A new description of trellis codes," *IEEE Trans. Inform. Theory,* vol.. IT-30, pp.784-791, Nov. 1984.

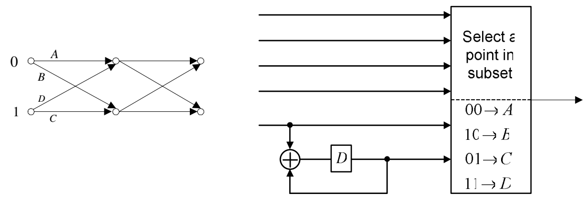


Fig. 1 Ungerboeck's 2-state TCM scheme: the trellis diagram and the encoder

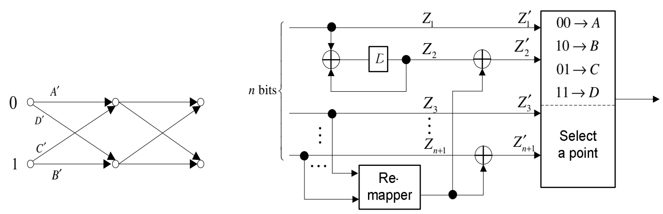


Fig. 2 Two-state TCM scheme with 4 partially overlapped signal sets

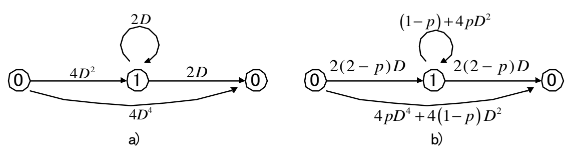
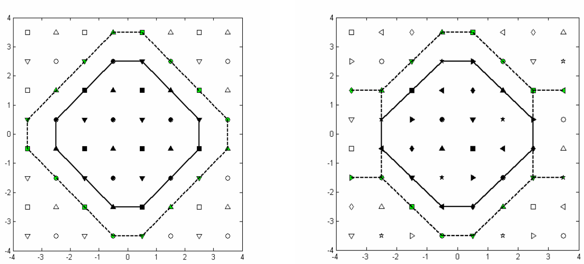


Fig. 3 Error state diagrams a) of Ungerboeck's code and b) of new code



a) b)

Fig. 4a. The 4-way partition: A - circle, C - triangle - down, B - square, D - triangle-up; Black-inner, White-outer; Dash line - 40-point signal set

Fig. 4b. The 8-way partition: A - circle, B - star, C - square, D - diamond, E - triangle-down, F - triangle-right, G - triangle-left, H - triangle-up; Black - inner, White - outer; Dash line - 40-point signal set

Table 1. Codes with 4-way partition of 64-QAM

# states	Ungerboeck, 64-QAM, 4-way partition		Overlapped, 40-QAM, 4-way partition	
	h^1	h^0	h^1	h^0
2	001	003	003	001
4	002	005	007	005
8	002	015	015	013
16	016	023	031	027
32	026	053	073	051
64	042	117	147	101
128	136	255	261	217

Table 2. Codes with 8-way partition of 64-QAM

# states	Ungerboeck, 64-QAM, 4-way partition			Overlapped, 40-QAM, 4-way partition		
	h^2	h^1	h^0	h^2	h^1	h^0
4	002	001	005	007	003	001
8	002	004	011	013	004	011
16	016	004	023	031	014	027
32	034	016	047	073	014	045
64	036	052	115	147	006	133

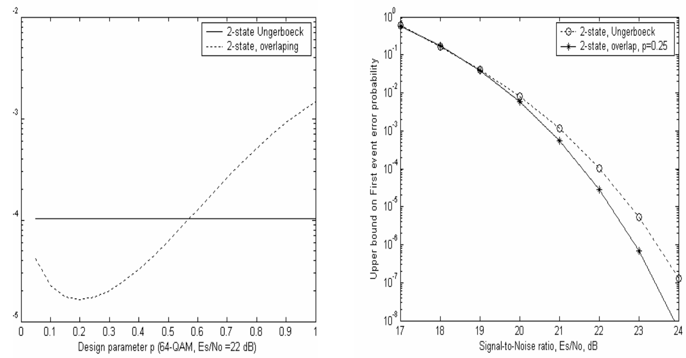


Fig. 5 Optimum value of the design parameter p and comparison of 2-state TCM codes with 64-QAM and 40-QAM

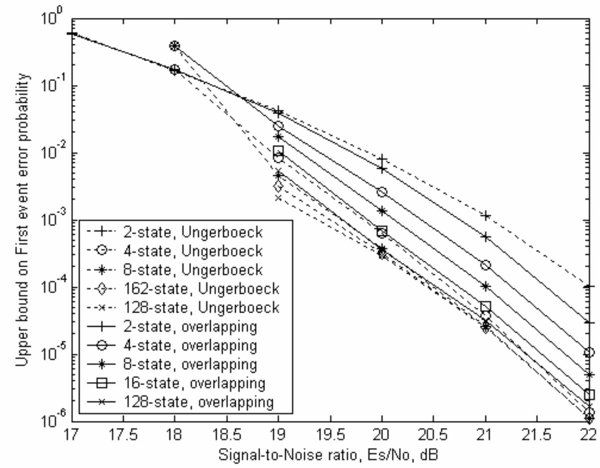


Fig. 6 TCM codes with the 4-way partition of the QAM constellation

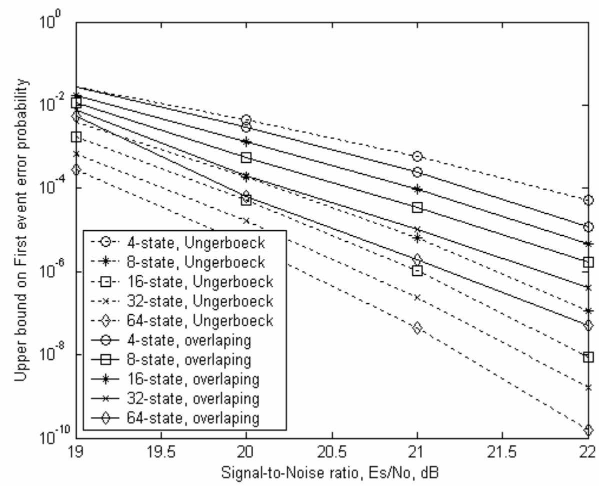


Fig. 7 TCM codes with the 8-way partition of the QAM constellation