TCM Without Constellation Expansion Penalty

Edit J. Kaminsky, James Ayo, and Kenneth V. Cartwright

Abstract: **We present a family of constant-amplitude constellations of even dimensions 8 and above. These constellations allow trellis coded modulation to be implemented without the usual penalty paid for constellation expansion. The new constellations are generated by concatenating either n QPSK points or n QPSK points rotated by 45 degrees, for any $n \geq 4$. Our constellations double the number of points available for transmission without decreasing the distance between points and without increasing the average or peak energies, introducing asymmetry, or increasing the modulation level. Effective gains of 2.65 dB with minimum complexity through 6.42 dB with moderate complexity are demonstrated using the 8D constellation.**

Index Terms: **Trellis coded modulation, constellation expansion, spherical codes, constant amplitude, multidimensional constellation.**

I. INTRODUCTION

We present a family of multi-dimensional constellations which are used in conjunction with a convolutional encoder to achieve simple trellis coded modulation (TCM) schemes with good asymptotic and effective gains. The constellations have even dimension $N = 2n$, with integer $n \geq 4$. The new constellations are generated by concatenating either n QPSK points, or n rotated QPSK points (each 2D point is rotated by 45 °). Thus, information transmission at 2b/s/Hz is performed over n consecutive time intervals. By allowing the rotated version of QPSK, the number of points in the resulting set is twice as large as the number of points of the original QPSK over n consecutive intervals, but the minimum Euclidean distance between points in the constellation, the modulation level, the peak energy, and the average energy remain identical to those in the constellation prior to expansion. The new constellations belong to the class of constant energy codes [1], and can be easily scaled to have unit energy and therefore be classified as spherical codes [1], [2].

Much work on TCM has been published since Ungerboek presented TCM as a combination of modulation and coding in [3]. For a TCM system of overall rate $k/(k+q)$, where k is the number of input bits into the TCM encoder, and $(k+q)$ is the number of output bits, the required coding redundancy is introduced with a convolutional encoder of rate $m/(m+q)$, where $m \leq k$. The resulting $k+q$ bits select, from within an expanded constellation of 2^{k+q} possible points, the one point to be transmitted.

Manuscript received April 27, 2001; approved for publication by J. J. Boutros, Division I Editor, November 15, 2001.

E. Kaminsky is with the Department of Electrical Engineering, University of New Orleans, New Orleans, LA 70148, U.S.A., e-mail: ejbourge@uno.edu.

J. Ayo is with the University of Tennesee, College of Business Administration, Knoxville, TN 37996, U.S.A., e-mail: jayo@ln.utk.edu.

K. Cartwright is with the College of The Bahamas, School of Technology, P.O. Box N4912, Nassau, NP, The Bahamas, e-mail: kvc@batelnet.bs.

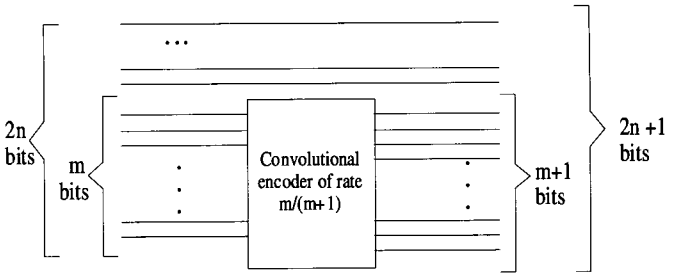


Fig. 1. TCM scheme of rate 2n/(2n+1)with one bit of redundancy introduced by a convolutional encoder of rate m/(m+1).

We will restrict our discussion to TCM schemes with $2n$ input bits and $2n+1$ output bits, therefore using convolutional encoders of rate $m/(m+1)$, introducing only 1 redundant bit per N-dimensional (ND) symbol. The expanded constellation must therefore contain twice as many points as the constellation that would be used for the uncoded case. A schematic representation of such a TCM scheme is shown in Fig. 1.

The disadvantage of constellation expansion required for TCM, more noticeable in low dimensions, is the reduced minimum squared Euclidean distance (MSED) between points for a given energy level, or the increase of modulation level and energy for a given MSED. For example, to transmit 2b/s/Hz with one bit of redundancy, the TCM system would use 8-PSK or 8 AM-PM, while the uncoded system would use QPSK [4]. To decrease the loss due to constellation expansion, higher-dimensional constellations are used [5], but there is still a loss associated with the constellations in prevalent use. A few authors have proposed systems which prevent the constellation expansion loss: Saha's codes [6] do not require any expansion of the original Q^2PSK signal set. Padovani and Wolf [7] combine frequency and phase modulation but their systems have a slightly different bandwidth. In [8] dual polarizations are used to double the dimension of the systems and allow for the introduction of redundancy for coding. The systems presented here utilize points in eight or more dimensions that do not have associated with them any expansion loss whatsoever; they also have constant amplitude, advantageous when nonlinear amplifiers are used. Our constellations also have the same modulation level and scheme as the uncoded system (basically QPSK over n consecutive intervals), and either are or are easily made invariant under rotations of 90, 180, and 270 degrees. The gain is achieved without sacrificing data rate or increasing the bandwidth of the system. The normalized redundancy [9] and the complexity of our codes are low, and in many cases the error coefficient is reduced below that of the uncoded system.

The use of the new constellations is exemplarized with convolutional encoders of low and moderate complexities for transmission of 2 b/s/Hz using quadrature modulation over four consecutive time intervals, but analysis may be easily extended to

1229-2370/02/$10.00

the other, higher dimensional constellations with lower normalized redundancy.

In Section II the new constellations are presented, followed in Section III by the set partitioning required for trellis encoding. In Section IV we discuss the various measures used to evaluate our TCM schemes. A summarized discussion of decoding and decoding complexity is presented in Section V. Results for the 8D constellation are presented in Section VI for encoders of various rates, including simulation results for the simplest codes. Comparisons to other TCM schemes are also given in Section VI. Concluding remarks and references follow.

II. THE NEW CONSTELLATIONS

The $2n$-dimensional ($2n$D) points in our new constellations are constructed by transmitting n consecutive 2D in-phase and quadrature (I, Q) pairs, with $n \geq 4$. If $I(nk+j)$ and $Q(nk+j)$, $j = 0, 1, 2, \cdots, n-1$, represent the I and Q data streams for the j^{th} time slot of symbol k, respectively, the k^{th} transmitted symbol may be written as $P_k = (p_0, p_1, \cdots, p_{n-1})$, where p_i is the 2D symbol given by (I, Q). For standard QPSK over n consecutive time intervals there are 4^n possible points P_k; for example, if 4 consecutive time intervals are used to generate 8D points, there are 256 possible values of P_k. However, in order to introduce 1 bit of redundancy for trellis coding, we need twice as many points, namely $2 \cdot 4^n$ points, or 512 points in 8D. To meet the requirement of constant envelope for all time, we require the energy normalized per 2D to remain at 2 as is the case with QPSK when $(\pm 1, \pm 1)$ are transmitted. We can satisfy both these requirements of constant envelope and doubling the constellation size without reducing the minimum squared Euclidean distance (MSED) between points. To do this we allow $2n$D symbols of the form $P_k r = (p_0 r, p_1 r, \cdots, p_{(n-1)} r)$, where $p_i r$ is p_i above rotated by 45 degrees. Hence, p_ir may be $(0, \mp\sqrt{2})$ or $(\pm\sqrt{2}, 0)$. Since $p_i r$ may take one of four values each, there are 4^n points of this type in n-dimensions. The new constellation for $2n$D is the union of $\{P_k\}$ and $\{P_k r\}$, yielding the $2 \cdot 4^n$ points required for TCM[1].

The MSED between any two points in the rotated $2n$D constellation $\{P_k r\}$ is 4, as it was with the original set $\{P_k\}$. The MSED between points in $\{P_k r\}$ and points in $\{P_k\}$ is $n + n(1 - \sqrt{2})^2$; the MSED is therefore 4 for all integers $n \geq 4$. In summary, the inclusion of the new 4^n points has not reduced the MSED. The new constellation of points is more densely packed in the sphere, utilizing the available $2n$D space better. Notice also that the peak and average energies of the new constellations are both $2n$, or 2 per 2D. Using these constellations, then, we can select from twice as many possible points with the same energy and at the same modulation level as uncoded QPSK over n consecutive time intervals; this is what makes our constellation more attractive than others which increase the modulation level to allow for redundancy. Moreover, it will be shown that decoding for this system is straightforward and of very low complexity even when the number of parallel transitions in the trellis is high.

[1]Notice that we could also use the expanded constellations without coding, to increase the transmission rate from 2 b/s/Hz to (2n+1)/n b/s/Hz.

Table 1. Properties of the new constellations.

n	N	Points	E	ρ	d_0^2	N(d_0)
4	8	512	8	0.25	4	8
5	10	2048	10	0.20	4	10
6	12	8192	12	0.17	4	12

Table 2. Distance distribution.

	8D		10D		12D	
Level i	d_i^2	N(d_i)	d_i^2	N(d_i)	d_i^2	N(d_i)
0	4	8	4	10	4	12
1	4.6863	16	5.8579	32	7.0294	64
2	8	28	8	45	8	66
3	10.3431	64	11.5147	160	12	220
4	12	56	12	120	12.6863	384
5	16	166	16	210	16	495

As an example, consider 8D transmission where $n = 4$ consecutive time slots are used for each symbol. The new constellation contains 512 points with MSED of 4, and both peak and average energies of 8. The points (1,1,-1,-1,1,1,-1,-1) and $(0, -\sqrt{2}, \sqrt{2}, 0, -\sqrt{2}, 0, 0, \sqrt{2})$ are two of the 512 possible 8D points.

Table 1 summarizes characteristics of the constellations for $n = 4, 5$, and 6. The column labeled N(d_0) indicates the number of points at MSED d_0^2; notice that the MSED of the trellis encoded (after partitioning) sets will be drastically larger than the MSED within the constellation, since those points at MSED are not assigned to the same trellis branches. The normalized redundancy [9], ρ, defined as the number of redundant bits per 2D, is also listed. In general, the new constellation using n consecutive 2D intervals yields dimension $N = 2n$, $2 \cdot 4^n$ for the number of points, an energy equal to the dimension, $E = N$, and normalized redundancy $\rho = 1/n$.

Table 2 presents the pre-partitioning distance distribution of the 8D, 10D, and 12D constellations up to SED of 16. The distance distribution is important because the minimum distance and the number of neighbours at that distance (error coefficient) determine the probability of error of a system using such set of points. We would like the minimun distance to be large, and the error coefficient to be small since, clearly, the most likely error is between closest points. The ideas behind constellation partitioning, discussed in the next section, are to increase the MSED between points in the partition, and minimize the error coefficient.

III. CONSTELLATION PARTITION

In order to have subsets to assign to the trellis branches of the convolutional encoder [4], we must partition the constellations so as to increase the MSED as much as possible, keep the number of nearest neighbours as small as possible, and maintain rotational invariance, if possible. Wei proposed a partitioning scheme for multidimensional constellations in [5]. This scheme must be slightly changed in order to partition our constellations to minimize the error coefficient $N(d)$,which is the number of nearest neighbours at distance d.

We use Wei's terminology, as presented in [5] whenever pos-

Table 3. 8D constellation partition.

Name	QPSK	Rotated QPSK	Process	(nr,r,tot)	MSED
2D Sublattices	A, B, C, D	Ar, Br, Cr, Dr	Original constellation	(4,4,8)	4
4D Types	$T_1,T_2,\cdots,T_{16}$	$T_1r, T_2r,\cdots, T_{16}r$	Concatenation of 2	(16,16,32)	4
4D Sublattices	$S_0,S_1,\cdots,S_7$	$S_0r, S_1r,\cdots, S_7r$	Union of 2	(8,8,16)	4
8D Types	$t_{00}, t_{01},\cdots, t_{77}$	$t_{00}r, t_{01}r,\cdots, t_{77}r$	Concatenation of 2	(64,64,128)	16
8D Subtype	$u_0, u_1,\cdots, u_{31}$	$u_0r, u_1r, \cdots, u_{31}r$	Union of 2	(32,32,64)	16
8D Sublattice	$s_0, s_1,\cdots, s_{15}$	$s_0r, s_1r,\cdots, s_{15}r$	Union of 2	(16,16,32)	16
8D Sublattice	e_0 and e_1	e_2 and e_3	Union of 8	(2,2,4)	8
8D Sublattice	$ss_0, ss_1, ss_2, ss_3, \cdots, ss_{15}$		Union of 2, mixed	(-,-,16)	10.34
8D Sublattice	$E_0, E_1, E_2, \cdots, E_7$		Union of 2, mixed	(-,-, 8)	8

Table 4. 8D sublattices s_i and subtypes u_i are generated by the union of the component 8D types indicated.

8D Sublattices s_i	8D Subtypes u_i
$s_0 = u_0 \cup u_1$	$u_0 = t_{00} \cup t_{11}$; $u_1 = t_{22} \cup t_{33}$
$s_1 = u_2 \cup u_3$	$u_2 = t_{01} \cup t_{10}$; $u_3 = t_{23} \cup t_{32}$
$s_2 = u_4 \cup u_5$	$u_4 = t_{02} \cup t_{13}$; $u_5 = t_{20} \cup t_{31}$
$s_3 = u_6 \cup u_7$	$u_6 = t_{03} \cup t_{12}$; $u_7 = t_{21} \cup t_{30}$
$s_4 = u_8 \cup u_9$	$u_8 = t_{44} \cup t_{55}$; $u_9 = t_{66} \cup t_{77}$
$s_5 = u_{10} \cup u_{11}$	$u_{10} = t_{45} \cup t_{54}$; $u_{11} = t_{67} \cup t_{76}$
$s_6 = u_{12} \cup u_{13}$	$u_{12} = t_{46} \cup t_{57}$; $u_{13} = t_{64} \cup t_{75}$
$s_7 = u_{14} \cup u_{15}$	$u_{14} = t_{47} \cup t_{56}$; $u_{15} = t_{65} \cup t_{74}$
$s_8 = u_{16} \cup u_{17}$	$u_{16} = t_{04} \cup t_{15}$; $u_{17} = t_{26} \cup t_{37}$
$s_9 = u_{18} \cup u_{19}$	$u_{18} = t_{05} \cup t_{14}$; $u_{19} = t_{27} \cup t_{36}$
$s_{10} = u_{20} \cup u_{21}$	$u_{20} = t_{06} \cup t_{17}$; $u_{21} = t_{24} \cup t_{35}$
$s_{11} = u_{22} \cup u_{23}$	$u_{22} = t_{07} \cup t_{16}$; $u_{23} = t_{25} \cup t_{34}$
$s_{12} = u_{24} \cup u_{25}$	$u_{24} = t_{40} \cup t_{51}$; $u_{25} = t_{62} \cup t_{73}$
$s_{13} = u_{26} \cup u_{27}$	$u_{26} = t_{41} \cup t_{50}$; $u_{27} = t_{63} \cup t_{72}$
$s_{14} = u_{28} \cup u_{29}$	$u_{28} = t_{42} \cup t_{53}$; $u_{29} = t_{60} \cup t_{71}$
$s_{15} = u_{30} \cup u_{31}$	$u_{30} = t_{43} \cup t_{52}$; $u_{31} = t_{61} \cup t_{70}$

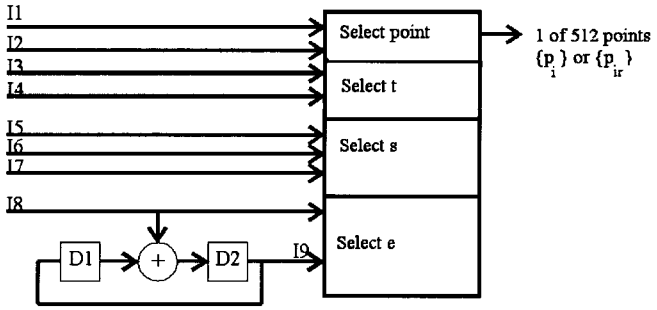


Fig. 2. TCM encoder showing the convolutional encoder of rate 1/2 and the mapping to the sets obtained from set partitioning.

sible, which we now summarize: A lattice [2] (the starting set) is partitioned into sublattices; only the bottom level is referred to by Wei as a sublattice. The process commences with constituent 2D lattices and ends with ND sublattices, which are the subsets assigned to trellis transitions. A sublattice may be further partitioned into types and subtypes with the same MSED. Each 2nD type is a concatenation of a pair of nD sublattices.

Only the partition for the 8D constellation is detailed here, and it is summarized in Table 3. The column labelled (nr, r, tot) lists the number of non-rotated QPSK subsets, the number of rotated-QPSK subsets, and the total number at the partition level indicated. Notice that the last two entries describe sublattices that are "mixed", in the sense that both rotated and non-rotated points are used in these subsets.

First we form four 2D sublattices for the QPSK and rotated QPSK 2D constellations: A =(1,1), B=(-1,-1), C=(1,-1), D=(-1,1), Ar=$(0,\sqrt{2})$, Br=$(0,-\sqrt{2})$, Cr=$(\sqrt{2},0)$, and Dr=$(-\sqrt{2},0)$. Next, we form the 32 4D types by concatenating a pair of points: T_1=(A,A), T_2=(A,B),$\cdots$, T_4=(A,D), T_5=(B,A),$\cdots$, T_{16}=(D,D), along with the corresponding T_1r through $T_{16}r$. Grouping, by union, a pair of these 4D types yields the 4D sublattices, eight for each of the two sub-constellations: $S_0=\{T_1 \cup T_6\}$, $S_1=\{T_{11} \cup T_{16}\}$, $S_2=\{T_2 \cup T_5\}$, $S_3=\{T_{12} \cup T_{15}\}$, $S_4=\{T_3 \cup T_8\}$, $S_5=\{T_{10} \cup T_{13}\}$, $S_6=\{T_4 \cup T_7\}$, $S_7=\{T_9 \cup T_{14}\}$, and the corresponding S_0r through S_7r. Now we form 8D types by concatenating a pair of these 4D sublattices, for a total of 128 types. We label them t_{00} through t_{77} and $t_{00}r$ through $t_{77}r$ where the two subscripts indicate the constituent S_i sublattices. The MSED between the 8D points within each type t_{ij} is 16 since at least 4 bits differ. For example, $t_{35} = \{S_3\ S_5\}=\{(T_{12}\ T_{10});\ (T_{12}\ T_{13});\ (T_{15}\ T_{10});\ (T_{15}\ T_{13})\}$ = {(1 -1 -1 1 1 -1 -1 -1); (1 -1 -1 1 -1 1 1 1); (-1 1 1 -1 1 -1 -1 -1); (-1 1 1 -1 -1 1 1 1)}.

Remember that the final partition subsets (sublattices) are to be assigned to the trellis transitons. The trellis will have as many parallel transitions as there are points within the subset. To reduce the number of subsets to be used for lower rate codes with few trellis states, we continue as follows: Form 8D subtypes, labelled u_0 through u_{31} and u_0r through $u_{31}r$, by union of the 8D types. These are detailed in Table 4, where only the 8D subtypes formed by the unrotated points are shown, but the rotated u_ir are composed in the same way using the rotated 8D types $t_{ij}r$. The MSED among the 8 points within each of the u_i sublattices is still 16.

Once again, we may group by union two of the u_i's, to reduce the number of subsets by a factor of two, while increasing the number of points within by a factor of two also. We call these new sublattices s_i, and detail their formation in Table 4. The MSED is still 16.

Further grouping must be used for lower rate codes, producing fewer subsets and, unfortunately, decreased MSED. More than one grouping path is now possible; which way is chosen depends on the number of subsets desired at the end of the partitioning process.

Table 5. Mixed 8D sublattices ss_i and E_i.

8D Sublattices E_i	8D Component Sublattices ss_i
$E_0 = ss_0 \cup ss_6$	$ss_0 = s_0 \cup s_5r$; $ss_1 = s_1 \cup s_4r$
$E_1 = ss_1 \cup ss_7$	$ss_2 = s_2 \cup s_0r$; $ss_3 = s_3 \cup s_1r$
$E_2 = ss_2 \cup ss_5$	$ss_4 = s_4 \cup s_6r$; $ss_5 = s_5 \cup s_7r$
$E_3 = ss_3 \cup ss_4$	$ss_6 = s_6 \cup s_3r$; $ss_7 = s_7 \cup s_2r$
$E_4 = ss_8 \cup ss_{13}$	$ss_8 = s_8 \cup s_{10}r$; $ss_9 = s_9 \cup s_{11}r$
$E_5 = ss_9 \cup ss_{12}$	$ss_{10} = s_{10} \cup s_8r$; $ss_{11} = s_{11} \cup s_9r$
$E_6 = ss_{10} \cup ss_{15}$	$ss_{12} = s_{12} \cup s_{14}r$; $ss_{13} = s_{13} \cup s_{15}r$
$E_7 = ss_{11} \cup ss_{14}$	$ss_{14} = s_{14} \cup s_{12}r$; $ss_{15} = s_{15} \cup s_{13}r$

Table 6. Partition of the 8D constellation.

Subset Labels	No. of subsets	No. pts. within	MSED within	No. at MSED
t_{00} - $t_{77}r$	128	4	16	2
u_0 - $u_{31}r$	64	8	16	6
s_0 - $s_{15}r$	32	16	16	14
ss_0 - $ss_{15}r$	16	32	10.3431	8
E_0 - E_7	8	64	8	4
e_0 - e_4	4	128	8	28

To group into 4 final subsets of 128 points each, with MSED=8, we create the sublattices $e_0 = \cup_{i=0}^{7} s_i$, $e_1 = \cup_{i=8}^{15} s_i$, $e_2 = \cup_{i=0}^{7} s_i r$, and $e_3 = \cup_{i=8}^{15} s_i r$. A block diagram of a simple system which uses this partition is shown in Fig. 2, along with references to the relevant partition levels.

If we desire 16 subsets, we combine sublattices s_i and $s_i r$, in what we've called "mixed sets" ss_i, as indicated in Table 5. There are 8 neighbours at MSED of 10.3431 within each of the above ss_i subsets. Joining pairs of these yields the 8 subsets to be used for a rate 2/3 encoder; these are made up by the union of two of the ss_i's, and are also shown in Table 5. The MSED in this case is 8, with only four points at such distance.

Details about all these groupings are given in Table 6.

IV. EVALUATION OF TCM SCHEMES

Many TCM schemes can be designed with our multidimensional constellations; we evaluate here only a few, all of them using the 8D constellation. The evaluation is based on the following standard measures: Asymptotic gain [4], effective gain [9], normalized redundancy [9], decoding complexity [10], modulation level and, ultimately, the probability of symbol error vs. signal to noise ratio (SNR) curves. We define SNR in terms of the symbol energy, E, and the variance of each component of the AWGN noise, σ^2, as

$$SNR = \frac{E}{N\sigma^2}. \tag{1}$$

The probability of symbol error bound is

$$P_{es} \geq N(d_{free}) Q\left(\frac{d_{free}}{2\sigma}\right), \tag{2}$$

where Q(·) is the Gaussian error-probability function and $N(d_{free})$ is the average number of pairs of trellis paths at distance d_{free}. The free distance, d_{free}, is determined by

$$d_{free} = \mathrm{Min}(\mathrm{d_{par}}, \mathrm{d_{min}}), \tag{3}$$

where d_{par} is the MED between parallel branches in the trellis, and d_{min} is the MED between paths of length longer than one branch. Asymptotic gain is given by

$$G_A = 10 \log_{10}\left(\frac{d_{free}^2}{d_u^2}\right), \tag{4}$$

where d_u^2 represents the MSED of the equivalent (same energy and same rate) uncoded system.

Forney [9] uses the rule of thumb that every factor of two increase in the error coefficient with respect to that of the uncoded system reduces the coding gain by about 0.2 dB. The loss, λ, is then given by:

$$\lambda = \frac{\log_{10}\left(\frac{N_{o,c}}{N_{o,u}}\right)}{\log_{10}(32)}, \tag{5}$$

where $N_{o,c}$ is the error coefficient of the coded system normalized per 2D, and $N_{o,u}$ is the corresponding normalized error coefficient for the uncoded system. This enables us to compute an effective coding gain, γ_{eff} :

$$\gamma_{eff} = G_A - \lambda. \tag{6}$$

Normalized redundancy, ρ, is the number of redundant bits per 2D symbol [9]. In our case, we use only one redundant bit in 8, 10, and 12D, yielding redundancies of 0.25, 0.20, and 0.17, respectively.

V. TCM DECODING

As far as the authors can determine, there is no single quantity used to characterize the TCM decoding complexity. TCM decoding is best understood if separated into two: Viterbi decoding and subset decoding. There is a complexity associated with each of these two phases; in both cases, we desire to have low complexity.

A. *Viterbi Decoding*

The Viterbi decoding process uses the standard Viterbi decoding algorithm [11], [12] to search the trellis and find the most likely path, given the received sequence of subsets. Real implementations truncate the decision length to about 5 times the constraint length of the code [12], therefore making a decision after a sequence of at least 5ν symbols have been received. Viterbi decoding for TCM is performed, basically, in the same way as it is done for convolutional codes.

For Viterbi decoding complexity evaluation, we use β as presented in [10]:

$$\beta = \log_2\left(\frac{2^{v+m}}{n}\right) = v + m - \log_2 n, \tag{7}$$

where m is the number of bits checked by the encoder, and v is the constraint length of the code. This quantity is an indication

of the decoding complexity of the Viterbi decoder, but provides no information about the complexity of the required subset decoding, i.e., the selection of the point within the selected subset.

B. Subset Decoding

Subset decoding chooses the point within the selected subset, that is closest to the received symbol. This, therefore, is the decoding of the parallel transitions in the trellis. The number of comparisons or distance computations is a partial indicator of the complexity of subset decoding. We emphasize that even for the simplest code presented here, the 8D TCM system with an encoder with R=1/2 and $v = 2$, which has 128 parallel branches per transition, (shown in Fig. 2), the complexity of the subset decoder is very low since no distance computations need to be performed in most cases, and only eight in the rest.

As mentioned when introducing the new constellations, no asymmetry or increase in the modulation level is introduced with our constellations; i.e., we are still basically dealing with QPSK, and the subset decoding is therefore quite simple. We will limit this discussion to decoding of the final 8D subsets e_0 through e_4; moreover, due to the similarity between all these subsets we can, without loss of generality, discuss only e_0. This is the decoding required for the system of Fig. 2.

Once the Viterbi decoding algorithm has selected the subset (e_0), a single point from within the selected subset must be chosen as the most likely transmitted symbol, selecting one of the 128 parellel transitions. Hard-limiter decoding (with a threshold of 0) of the received signal is performed independently in each of the 8 dimensions. A table look-up operation is then performed to determine whether the hard-limited symbol belongs to the selected subset; if it does, it is chosen as the decoded symbol without any distance computations performed. If it does not belong to the selected subset, then 8 distance computations must be performed. Based on the way the partition was done, if the thresholded symbol does not belong to the subset of interest, a symbol that differs from it in only one position will. In 8D, there are eight of these symbols, and one of these will be closest to the received symbol. The received symbol, then, is compared to these 8 candidate symbols, and the closest is selected as the decoded symbol. This decoding method gives exactly the same performance as the more computationally intensive method of comparing the received symbol to the 128 points within the subset selected by the Viterbi decoder.

VI. RESULTS

We present in this section the evaluation results of several schemes using the 8D constellation together with convolutional encoders of various rates. Higher gains are achieved as the complexity increases. Similar schemes can be developed for the higher dimensional constellations, but results for dimension higher than 8 are not discussed in the current paper. We separate the discussion and presentation of results according to the system's asymptotic gain. All the systems evaluated use the 8D constellation and the partitions of Table 3. We follow the discussion of our codes with a comparison to other TCM schemes used to transmit 2 b/s/Hz.

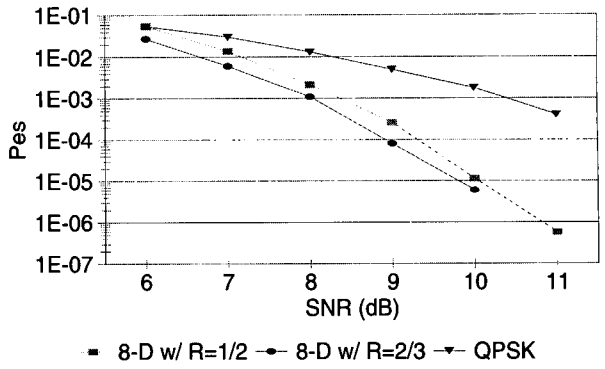


Fig. 3. Probability of symbol error, Pes, versus symbol SNR, in dB, of coded systems with new 8D constellation. Results are shown for encoders of rate 1/2 and 2/3, with 4 and 8 trellis states, respectively, as well as for the uncoded QPSK reference system.

A. 8D TCM Systems Yielding G_A=3.01 dB

The convolutional encoder of rate 1/2, shown in Fig. 2, yields an extremely simple system which uses our 8D constellation. The 512 points are partitioned into the 4 subsets e_0 through e_3 of 128 points each. The system's complexity, as defined by (7), is $\beta = 2 + 1 - \log_2 4 = 1$ since only one bit is checked by the encoder, and the constraint length is $v = 2$, with $n = 4$.

The free distance is determined by the parallel transitions, and is $d^2_{free} = 8$, the MSED among points in any e_i sublattice. The MSED among uncoded points (standard QPSK) is 4, yielding an asymptotic gain of $G_A = 10 \log_{10}(2) = 3.01$ dB. The losses due to the error coefficient are determined by $N_{o,c} = 7$ (28 paths in 8D) and $N_{o,u} = 2$, and are $\lambda = 0.36$ dB, yielding an effective gain of $\gamma_{eff} = 2.65$ dB. The redundancy is 1 bit per 8D, or $\rho = 0.25$ normalized per 2D.

By increasing the rate of the convolutional encoder to 2/3, with 8 states in the trellis, and 3 memory elements, we may use the partition sets E_0-E_7. The free MSED, d^2_{free}, is still 8, but the error coefficient is reduced to $N_{o,c} = 1$ (4 parallel paths in 8D) yielding additional gains of 0.2 dB. The effective gain is therefore $\gamma_{eff} = 3.21$ dB. The Viterbi decoding complexity is now $\beta = 3 + 2 - 2 = 3$, with little subset decoding complexity, and a redundancy of $\rho = 0.25$.

Fig. 3 presents simulation results in the form of probability of symbol error, Pes, vs. SNR, in dB, for the system of Fig. 2, using a Viterbi decoder truncated to 12 symbols. We define SNR as the ratio of total energy E to total noise energy $N\sigma^2$. The performance of the system with the convolutional encoder with R=2/3, also shown in Fig. 3, is just slightly better than that of the coded 8-PSK system presented in [4], which uses the same 8-state encoder; this improvement is due to the reduction in the error coefficient when our constellation is used. The performance of the baseline uncoded QPSK over 4 consecutive intervals is also shown in that plot.

B. 8D TCM System Yielding G_A=4.13 dB

If we partition the 8D constellation into the 16 ss_i sublattices of 32 points each, discussed in Section III, and use an encoder of rate 3/4 to introduce one bit of redundancy every 4

Table 7. Codes with 8D constellation.

Rate	β	d^2_{free}	$N(d_{free})$	$N_{o,c}$	G_A	λ	γ_{eff}
1/2	1	8	28	7	3.01	0.36	2.65
2/3	3	8	4	1	3.01	-0.20	3.21
3/4	5	10.3431	8	2	4.13	0.00	4.13
4/5	7	16	14	3.5	6.02	0.16	5.86
5/6	9	16	6	1.5	6.02	-0.08	6.10
6/7	11	16	2	0.5	6.02	-0.40	6.42

signalling intervals, we can increase the MSED to 10.3431. We therefore have G_A =4.13 dB with $N_{o,c} = 8/4$, which yields $\gamma_{eff} = G_A$ =4.13 dB. Again, there is no loss due to constellation expansion, and no loss due to nearest neighbours since the error coefficient remains the same as for the uncoded system. The complexity of the Viterbi decoder for this scheme is $\beta = 5$. Subset decoding is similar to that presented in Section V-B, except that if the hard limited symbol is not in the chosen subset, the subset it belongs to must be identified through lookup tables before (a small number) of distances are computed.

C. Higher Gain Codes with 8D Constellation

By increasing the rate and the number of states of the convolutional encoder, we can increase the MSED between parallel transitions to 16. Clearly, then, the asymptotic gain is increased to $G_A = 6.02$ dB. The error coefficient decreases as the rate and number of states of the encoder increases, but the complexity of the Viterbi decoding (not the subset decoding) increases. In all cases the redundancy is still $\rho = 0.25$.

Table 7 summarizes details for the 8D codes just discussed; it includes the complexity β, the free squared distance d^2_{free}, the error coefficient in 8D and normalized to 2D, $N(d_{free})$ and $N_{o,c}$,respectively, and the associated loss or gain λ, as well as the asymptotic and effective gains of the system, in dB, G_A and γ_{eff}. TCM systems with convolutional encoders of rate 5/6 and above are not used as often due to the increased complexity of the Viterbi decoding process, but are included for completeness and to demonstrate the additional gains arising from the reduction in the error coefficient.

D. Comparison to Other TCM Systems

We have gathered some comparison data in Table 8, where we only show results for encoders of rate 2/3 and 3/4. We include two of our codes and several others presented in the open literature, for the same transmission rate of 2 b/s/Hz. We concentrate on codes using multidimensional constellations with moderate complexities, represented by ν, the number of memory elements. The comparison is by no means comprehensive, but should confirm the benefits of our codes, even though our gains are not always greater than those of other TCM schemes. We, again, stress that the advantage of using our constellations is the simplicity of implementation and decoding, and the lack of constellation expansion penalty without an increase in modulation level.

From the entries in Table 8, we see that our codes are competitive, in particular when effective gain and complexity are used to compare codes. Notice that Divsalar's rate-3/4 code [13] has the highest asymptotic gain, but that asymmetric constellations

Table 8. Comparison of TCM systems with convolutional encoders of rate 2/3 and 3/4.

Rate	ν	N	ρ	G_A	γ_{eff}	Ref.
2/3	2	8	1/4	3.01	3.21	Here
2/3	2	8	1/4	3.01	2.89	[10]
2/3	4	8	1/4	3.01	2.89	[10]
2/3	4	4	1	4.15	NA	[13]
3/4	3	8	1/4	4.13	4.13	Here
3/4	3	8	5/4	5.27	3.75	[9]
3/4	6	8	5/4	5.27	4.49	[9]
3/4	3	8	1/4	4.02	NA	[9]
3/4	6	8	1/4	5.27	4.49	[9]
3/4	3	8	1/4	3.01	3.01	[10]

are used. For the rate 3/4 codes, some of Forney's codes [9] achieve higher gains, but at the expense of more redundancy and/or higher complexity of the trellis.

VII. CONCLUSIONS

We have presented trellis coded modulation using novel spherical multidimensional constellations which are twice the size of the constellation used for uncoded modulation, have the same average and peak energies, same modulation level, and same minimum distances between points. Asymptotic gains between 3.01 and 6.02 dB are easily achieved, with no loss due to constellation expansion. Additional gain is achieved in some cases by a reduction in the error coefficient under that of the uncoded system. The systems either are or may be made invariant under rotations of 90, 180, and 270 degrees with the use of differential encoders. Decoding is performed with a Viterbi decoder and simple subset decoding using hard limiters and lookup tables.

REFERENCES

[1] T. Ericson and V. Zinoviev, "Spherical codes generated by binary partitions of symmetric pointsets," *IEEE Trans. Inform. Theory*, vol. 41, no. 1, pp. 107–129, Jan. 1995.

[2] J. H. Conway and N. J. Sloane, *Sphere packings, lattices and groups*, Springer-Verlag, NY, 2nd. edition, 1993.

[3] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inform. Theory*, vol. 28, no. 1, pp. 55–66, Jan. 1982.

[4] G. Ungerboeck, "Trellis-coded modulation with redundant signal sets, part I: Introduction," *IEEE Commun. Mag.*, vol. 25, no. 2, pp. 5–11, Feb. 1987.

[5] L. Wei, "Trellis-coded modulation with multidimensional constellations," *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 483–501, July 1987.

[6] D. Saha, "Channel coding with quadrature-quadrature phase shift-keying (Q^2PSK) signals," *IEEE Trans. Commun.*, vol. 38, pp. 409–417, Apr. 1990.

[7] R. Padovani and J. Wolf, "Coded phase/frequency modulation," *IEEE Trans. Commun.*, vol. COM-34, pp. 446–453, May 1986.

[8] E. J. Kaminsky, *Trellis coding and adaptive estimation in dually polarized systems*, Ph.D. thesis, Tulane University, New Orleans, LA, June 1991.

[9] G. Forney, "Coset codes. part I: Introduction and geometrical classification," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1123–1151, Sept. 1988.

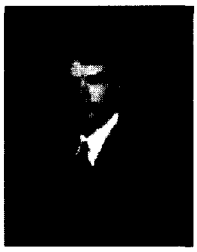
[10] S. Pietrobon *et al.*, "Trellis-coded multidimensional phase modulation," *IEEE Trans. Inform. Theory*, vol. 36, pp. 63–89, Jan. 1990.

[11] A. Viterbi, "Convolution codes and their performance in communication systems," *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 260–269, Apr. 1971.

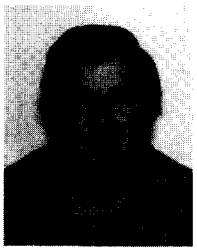
[12] G. Forney, "Convolution codes II: Maximum likelihood decoding," *Information and Control*, vol. 25, pp. 222–266, July 1974.

[13] D. Divsalar, M. Simon, and J. Yuen, "Trellis coding with asymmetric modulation," *IEEE Trans. Commun.*, vol. COM-35, no. 2, pp. 130–141, Feb. 1987.

Edit J. Kaminsky received her Ph.D and M.S. Degrees in Electrical Engineering from Tulane University, New Orleans, LA, USA, in 1991 and 1987, respectively. Her 1986 BSEE degree is from Universidad Autónoma de Centro América, Costa Rica. She is currently an Associate Professor of Electrical Engineering at the University of New Orleans, where she has taught and performed research since 1995. Prior to her academic career, she worked for Sverdrup Technology and Lockheed at Stennis Space Center, MS. Her research interests are in digital communications, coding, neural networks, and fuzzy systems. She is a member of IEEE, ASEE, SWE, Tau Beta Pi, and Eta Kappa Nu.



James Ayo is a graduate student at the University of Tennessee, Knoxville. His research interests include telecommunications and the application of technology. He received his B.S. in electrical engineering from Louisiana Tech University, Ruston, in 1997, and his M.S. in electrical engineering from the University of New Orleans, New Orleans, in 2000. Currently, he is working towards his MBA in Finance and Operations.



Kenneth V. Cartwright was born in Nassau, N. P., Bahamas in 1953. He received the B.E.Sc, M.S., and PhD degrees in electrical engineering from the University of Western Ontario in 1978, and Tulane University in 1987 and 1990, respectively. From 1978 to 1983, he was the service manager for a consumer electronics firm. He joined The College of The Bahamas in 1983 as an assistant lecturer and was promoted to lecturer in 1986. From 1985-1990, he was awarded study leave to pursue graduate work at Tulane University, where he was also an instructor in the Electrical Engineering Department for the 1988-1989 academic year. While at Tulane, he was awarded the Seto Award for Outstanding Graduate Student (1989 - 1990). He returned to The College of The Bahamas in 1990, where he received awards for Outstanding Performance for the academic years 1992-1993 and 1993-1994. Presently, he is a senior lecturer in the School of Technology and the Coordinator for the Electrical Engineering Department. His research interests include digital signal processing, analog and digital communication systems, and analog electronics. Some of this research is published in IEEE Transactions on Education, IEEE Signal Processing Letters, IEEE Communication Letters, IEEE Transactions on Communications, and conference proceedings. He is a member of Eta Kappa Nu, Tau Beta Pi, numerous IEEE societies, and ASEE.