

$y) \sin(x + y) + (x - y)(\sin y - \sin x)$ is always ≥ 0 for x and y in this region. Thus $f(x, y) \leq 4$ and $d_B^2 \leq 4$ for all $h_1, h_2 \leq 1.5$.

REFERENCES

- [1] H. Miyakawa, H. Harashima, and Y. Tanaka, "A new digital modulation scheme, multimode binary CPFSK," in Proc. Third Int. Conf. on Digital Satellite Commun., Nov. 1975, 11-13, Kyoto, Japan, pp. 105-112.
- [2] J. B. Anderson and D. P. Taylor, "A bandwidth-efficient class of signal space codes," *IEEE Trans. Inform. Theory*, vol. IT-24, no. 6, pp. 703-712, Nov. 1978.
- [3] J. B. Anderson and R. de Buda, "Better phase-modulation error performance using trellis phase codes," *Electron. Lett.*, vol. 12, no. 22, pp. 587-588, Oct. 28, 1976.
- [4] J. B. Anderson, "Simulated error performance of multi- h phase codes," in *Conf. Rec. Int. Conf. on Commun.*, Seattle, WA, June 1980, pp. 26.4.1-26.4.5.
- [5] J. B. Anderson, D. P. Taylor, and A. T. Lereim, "A class of trellis phase modulation codes for coding without bandwidth expansion," in *Conf. Rec. Int. Conf. on Commun.*, Toronto, ON, Canada, June 1978, pp. 50.3.1-50.3.5.
- [6] A. T. Lereim, "Spectral properties of multi- h phase codes," M. Eng. Thesis, McMaster Univ., Techn. Rep. no. CRL-57, July 1978, Communications Research Laboratory, Hamilton, ON, Canada.
- [7] J. B. Anderson, C. E. Sundberg, T. Aulin, and N. Rydbeck, "Power-bandwidth performance of smoothed phase modulation codes," *IEEE Trans. Commun.*, vol. COM-29, no. 3, pp. 187-195, Mar. 1981.
- [8] T. Aulin and C. E. Sundberg, "Continuous phase modulation (CPM), part I, full response signalling," *IEEE Trans. Commun.*, vol. COM-29, no. 3, pp. 196-209, Mar. 1981.
- [9] T. Aulin, N. Rydbeck, and C. E. Sundberg, "Continuous phase modulation (CPM)—Part II: Partial response signalling," *IEEE Trans. Commun.*, vol. COM-29, no. 3, pp. 210-225, Mar. 1981.
- [10] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. New York: Wiley, 1965.
- [11] W. P. Osborne and M. B. Luntz "Coherent and Noncoherent detection of CPFSK," *IEEE Trans. Commun.*, vol. COM-22, pp. 1023-1036, Aug. 1974.
- [12] R. de Buda, "About optimal properties of fast frequency-shift keying," *IEEE Trans. Commun.*, vol. COM-22, pp. 1974-1975, Oct. 1974.
- [13] T. A. Schonhoff, "Symbol error probabilities for M -ary CPFSK: coherent and noncoherent detection," *IEEE Trans. Commun.*, vol. COM-24, no. 6, pp. 644-652, June 1976.
- [14] T. Aulin, N. Rydbeck, and C. E. Sundberg, "Performance of constant envelope M -ary digital FM systems and their implementation," in *Conf. Rec. Nat. Telecommun. Conf.*, Washington, DC, Nov. 1979, pp. 55.1.1-55.1.6.
- [15] T. Aulin, N. Rydbeck, and C. E. Sundberg, "Transmitter and receiver structures for M -ary partial response FM," in *Proc. 1980 Int. Zürich Seminar on Digital Commun.*, Mar. 1980, pp. A.2.1-A.2.6.
- [16] T. Aulin, "Viterbi detection of continuous phase modulated signals," in *Conf. Rec. Nat. Telecommun. Conf.*, Houston, TX, Nov. 1980, pp. 14.2.1-14.2.7.
- [17] T. Aulin and C. E. Sundberg, "Minimum Euclidean distance and power spectrum for a class of smoothed phase modulation codes with constant envelope," University of Lund, Nov. 1980 (submitted to *IEEE Trans. Commun.*).
- [18] S. G. Wilson, J. H. Highfill III, and C. D. Hsu, "Error bounds for multi- h phase codes," University of Virginia, Charlottesville (submitted to *IEEE Trans. Inform. Theory*).
- [19] J. K. Omura and D. Jackson, "Cutoff rates for channels using bandwidth efficient modulations," in *Conf. Rec. Nat. Telecommun. Conf.*, Houston, TX, Nov. 1980, pp. 14.1.1-14.1.11.
- [20] T. Aulin, N. Rydbeck, and C. E. Sundberg, "Minimum distance and spectrum for M -ary multi- h constant envelope digital FM signals," Tech. Rep. TR-129, May 1979, Telecommun. Theory, Univ. Lund, Sweden.
- [21] T. Aulin and C. E. Sundberg, "Further results on M -ary multi- h constant envelope digital FM signals," Tech. Rep. TR-139, Apr. 1980, Telecommun. Theory, Univ. Lund, Sweden.

Channel Coding with Multilevel/Phase Signals

GOTTFRIED UNGERBOECK, MEMBER, IEEE

Abstract—A coding technique is described which improves error performance of synchronous data links without sacrificing data rate or requiring more bandwidth. This is achieved by channel coding with expanded sets of multilevel/phase signals in a manner which increases free Euclidean distance. Soft maximum-likelihood (ML) decoding using the Viterbi algorithm is assumed. Following a discussion of channel capacity, simple hand-designed trellis codes are presented for 8 phase-shift keying (PSK) and 16 quadrature amplitude-shift keying (QASK) modulation. These simple codes achieve coding gains in the order of 3-4 dB. It is then shown that the codes can be interpreted as binary convolutional codes with a mapping of coded bits into channel signals, which we call "mapping by set partitioning." Based on a new distance measure between binary code sequences which efficiently lower-bounds the Euclidean distance between

the corresponding channel signal sequences, a search procedure for more powerful codes is developed. Codes with coding gains up to 6 dB are obtained for a variety of multilevel/phase modulation schemes. Simulation results are presented and an example of carrier-phase tracking is discussed.

I. INTRODUCTION

IN CHANNEL CODING of the "algebraic" coding type, one is traditionally concerned with a discrete channel provided by some given modulation and hard-quantizing demodulation technique. Usually, inputs and outputs of the channel are binary. The ability to detect and/or correct errors can only be provided by the additional transmission of redundant bits, and thus by lowering the effective information rate per transmission bandwidth.

Manuscript received May 24, 1977; revised January 7, 1981.

The author is with the IBM Zurich Research Laboratory, 8803 Rüschlikon, Switzerland.

In addition, hard amplitude or phase decisions made in the demodulator prior to final decoding cause an irreversible loss of information. In the binary case, this amounts to loss equivalent to approximately 2 dB in signal-to-noise ratio (SNR).

In this paper, we take the viewpoint of "probabilistic" coding and decoding and regard channel coding and modulation as an entity [1]. Comparisons are strictly made on the basis of equal data rate and bandwidth. A possibility for redundant coding can therefore only be created by using larger sets of channel signals than required for nonredundant (uncoded) transmission. Maximum-likelihood (ML) soft decoding of the unquantized demodulator outputs is exclusively assumed, thus avoiding loss of information prior to final decoding. This implies that codes for multilevel/phase signals should be designed to achieve maximum free Euclidean distance rather than Hamming distance. For coded 2-amplitude modulation (AM) and 4-phase-shift keying (PSK) modulation, this was never a problem because in this case, binary Hamming distance (HD) and Euclidean distance (ED) are equivalent. The situation is different, however, if signal sets are expanded beyond two signals in one modulation dimension.

The investigations leading to this paper started some time ago with the heuristic design of simple trellis codes for 8-PSK modulation conveying two bits of information per modulation interval. When soft ML-decoded by the well-known Viterbi algorithm (VA) [2], coding gains of 3–4 dB were found over conventional uncoded 4-PSK modulation. The investigations were then extended to other modulation forms. First results were presented in [3]. A much better understanding of the subject was later obtained by interpreting the hand-designed codes as binary convolutional codes with a mapping of coded bits into multilevel/phase channel signals called "mapping by set partitioning."

Related work was reported in [4]–[8]. The approaches taken in [6]–[8] aim at constant-envelope modulation which is in contrast to the present paper. In [9], a comparison of various "bandwidth-efficient" modulation techniques by computer simulation is presented which includes codes of this paper.

In Section II, we investigate the potential gains in terms of channel capacity obtained by introducing more signal levels and/or phases. The results are similar to those of Wozencraft and Jacobs [10] on the exponential bound parameter R_0 , and suggest that for coded modulation, it will be sufficient to use twice the number of channel signals than for uncoded modulation. In Section III, heuristically designed trellis codes for coded 8-PSK and 16-QASK modulation are presented and the concept of mapping by set partitioning is introduced. The codes are interpreted in Section IV as binary convolutional codes of rate $R = m/(m+1)$ with the above mapping of coded bits into channel signals. Preference is given to realizations in the form of systematic encoders with feedback. The mapping rule allows the definition of a new distance measure that can easily be applied to binary code sequences and efficiently lower-bounds the ED between the correspond-

ing channel-signal sequences. Based on this distance measure, a search procedure for more powerful codes is developed in Section V, and codes with coding gains up to 6 dB are obtained for a larger variety of coded one- and two-dimensional modulation schemes. In Section VI simulation results are presented. Finally in Section VII, the problem of carrier-phase tracking, which can play an important role in the practical application of coded two-dimensional modulation schemes, is discussed for one specific case of coded 8-PSK modulation.

II. CHANNEL CAPACITY OF MULTILEVEL/PHASE MODULATION CHANNELS

Before addressing the code-design problem with expanded channel-signal sets, it is appropriate to examine in terms of channel capacity the limits to performance gains which may thereby be achieved. Because of our intended use of soft ML-decoding in the receiver, we must study modulation channels with discrete-valued multilevel/phase input and continuous-valued output. One- and two-dimensional modulation is considered and intersymbol interference-free signaling over bandlimited channels with additive white Gaussian noise (AWGN) is assumed. With perfect timing and carrier-phase synchronization, we sample at time $nT + \tau_s$ where T is the modulation interval and τ_s the appropriate sampling phase. The output of the modulation channel becomes

$$z_n = a_n + w_n, \quad (1)$$

where a_n denotes a real- or complex-valued discrete channel signal transmitted at modulation time nT , and w_n is an independent normally distributed noise sample with zero mean and variance σ^2 along each dimension. The average SNR is defined as

$$\begin{aligned} \text{SNR} &= \frac{E\{|a_n|^2\}}{E\{|w_n|^2\}} \\ &= \begin{cases} E\{|a_n|^2\}/\sigma^2 \cdots \text{(a) one-dimensional modulation} \\ E\{|a_n|^2\}/2\sigma^2 \cdots \text{(b) two-dimensional modulation} \end{cases} \end{aligned} \quad (2)$$

Fig. 1 illustrates the channel-signal sets considered in this paper. Normalized average signal power $E\{|a_n|^2\} = 1$ is assumed.

Extension of the well-known formula for the capacity of a discrete memoryless channel [11] to the case of continuous-valued output yields

$$C = \max_{Q(0) \cdots Q(N-1)} \sum_{k=0}^{N-1} Q(k) \int_{-\infty}^{+\infty} p(z/a^k) \cdot \log_2 \left\{ \frac{p(z/a^k)}{\sum_{i=0}^{N-1} Q(i)p(z/a^i)} \right\} dz \quad (3)$$

in bit/T. N is the number of discrete channel input signals

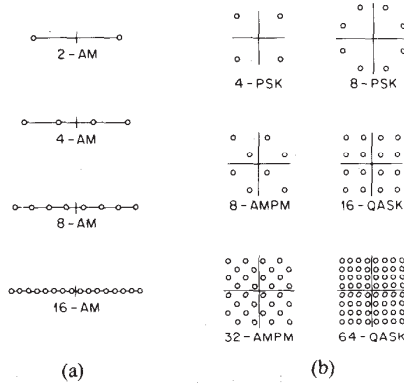


Fig. 1. Channel-signal sets considered in this paper. (a) One-dimensional modulation. (b) Two-dimensional modulation.

$\{a^0 \dots a^{N-1}\}$ and $Q(k)$ denotes the *a priori* probability associated with a^k . Because of AWGN, we can substitute in (3)

$$P(z/a^k) = \exp \left[-|z - a^k|^2 / 2\sigma^2 \right]$$

$$\begin{cases} (2\pi\sigma^2)^{-1/2} & \dots & (a) \\ (2\pi\sigma^2)^{-1} & \dots & (b) \end{cases} \quad (4)$$

With the further assumption that only codes with equiprobable occurrence of channel input signals are of interest, the maximization over the $Q(k)$ in (3) can be omitted. Equation (3) can now be written in the form

$$C_{Q(k)=1/N}^* = \log_2(N) - \frac{1}{N} \sum_{k=0}^{N-1} E \left\{ \log_2 \sum_{i=0}^{N-1} \exp \left[-\frac{|a^k + w - a^i|^2 - |w|^2}{2\sigma^2} \right] \right\} \quad (5)$$

In (5) we have integration replaced by expectation over the normally distributed noise variable w which is real with variance σ^2 for (a), and complex with variance $2\sigma^2$ for (b). Using a Gaussian random number generator, C^* has been evaluated by Monte Carlo averaging of (5). In Figs. 2(a) and 2(b), C^* is plotted as a function of SNR for the signal sets depicted in Fig. 1. The value of SNR at which in uncoded transmission symbol-error probability $Pr(e) = 10^{-5}$ is achieved is also indicated.

In order to interpret Figs. 2(a) and 2(b), we consider as an example transmission of 2 bit/T by uncoded 4-PSK modulation where $Pr(e) = 10^{-5}$ occurs at SNR = 12.9 dB. If the number of channel signals is doubled, e.g., by choosing 8-PSK modulation, error-free transmission of 2 bit/T is theoretically possible already at SNR = 5.9 dB (assuming unlimited coding and decoding effort). Beyond this—with no constraint on the number of signal levels/phases except average signal power—only 1.2 dB can further be gained. Similar proportions hold for the other modulation schemes. It can be concluded that by doubling the number of channel signals, almost all is gained in terms of channel capacity that is achievable by

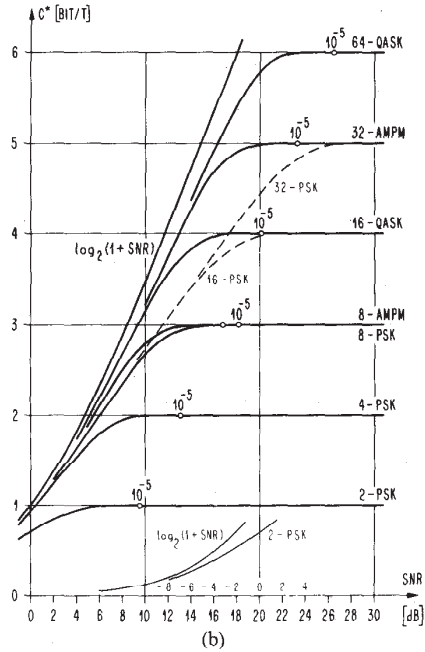
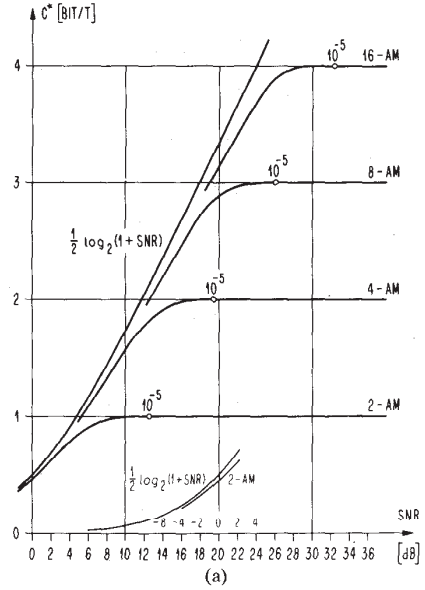


Fig. 2. Channel capacity C^* of bandlimited AWGN channels with discrete-valued input and continuous-valued output. a) One-dimensional modulation. b) Two-dimensional modulation.

signal-set expansion if at given SNR satisfactory error performance can no longer be achieved by uncoded modulation.

III. SIMPLE TRELLIS CODES

Coding gains can be realized either by block coding or by state-oriented trellis coding. There exists also the possibility of concatenation, e.g., by assigning short block-code words to state transitions in a trellis structure. Note that the choice of a signal set for two-dimensional modulation

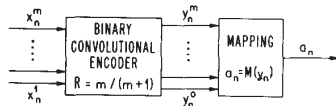


Fig. 3. Multilevel/phase encoder structure.

corresponds already to a simple form of block coding. In this paper however we do not further pursue the block coding aspect because the richer structure and omission of block boundaries together with the availability of Viterbi ML-decoding algorithm make trellis codes appear to us more attractive for the present coding problem.

Following the arguments of Section II, in order to improve error performance, m bit/ T must be transmitted in redundantly coded form by a set of 2^{m+1} channel signals. The coding expert will easily conclude that this may be accomplished by expanding the binary data sequence by suitable convolutional encoding with rate $R = m/(m+1)$, and subsequent mapping of groups of $m+1$ bits into the larger set of channel signals. The encoder structure is shown in Fig. 3. With $d(a_n, a'_n)$ denoting the ED between channel signals a_n and a'_n , the encoder should be designed to achieve maximum free ED:

$$d_{\text{free}} = \min_{\{a_n\} \neq \{a'_n\}} \left[\sum_n d^2(a_n, a'_n) \right]^{1/2} \quad (6)$$

between all pairs of channel-signal sequences $\{a_n\}$ and $\{a'_n\}$ which the encoder can produce. If soft ML-decoding is applied, the error-event probability will approach asymptotically at high SNR the lower bound [2]

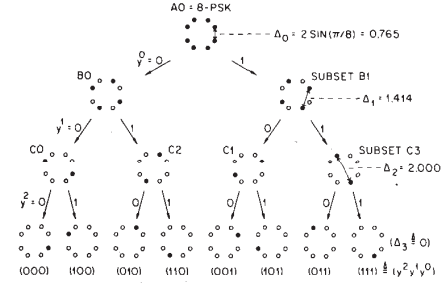
$$\Pr(e) \approx N(d_{\text{free}}) \cdot Q(d_{\text{free}}/2\sigma). \quad (7)$$

Here $N(d_{\text{free}})$ denotes the (average) multiplicity of error events with distance d_{free} , and $Q(\cdot)$ is the Gaussian error-probability function.

For transmission at 2 bit/ T by coded 8-PSK modulation it has been suggested [4], [5] to use known $R = 2/3$ binary convolutional codes with maximum free HD for given constraint length [12], and Gray coding as a mapping function. Yet there are problems with this approach. Firstly, the Gray code mapping does not monotonically translate larger HD into larger ED and secondly, permutations of the binary outputs of the convolutional encoder will have an unknown, perhaps significant influence on the ED structure of the resulting 8-PSK codes. Neither does the approach seem to be extendable to all modulation forms of Fig. 1.

We will therefore pursue a different design method which aims more directly at maximizing free ED. The approach is based on a mapping rule called "mapping by set partitioning." This mapping follows from successive partitioning of a channel-signal set into subsets with increasing minimum distances $\Delta_0 < \Delta_1 < \Delta_2 \dots$ between the signals of these subsets. The concept is illustrated in Figs. 4 and 5 for 8-PSK and 16-QASK modulation, respectively, and is applicable to all modulation forms of Fig. 1.

Before addressing the systematic search for convolutional codes for the encoder suggested by Fig. 3, we discuss

Fig. 4. Partitioning of 8-PSK channel signals into subsets with increasing minimum subset distances ($\Delta_0 < \Delta_1 < \Delta_2$; $E\{|a_n^2|\} = 1$).

in the remainder of this section the heuristic construction of simple but already very effective codes. This does not require knowledge of convolutional codes and will establish an intuitive basis for the development of more powerful codes later in the paper. Work leading to this paper progressed also in this order.

We regard an encoder simply as a finite-state machine with a given number of states and specified state transitions. If m bits are to be encoded per modulation interval T , there must be 2^m possible transitions from each state to a successor state. More than one transition may occur between pairs of states, and for obvious reasons only regular structures are of interest. After selecting a suitable trellis state-transition diagram, the remaining task consists of assigning channel signals from an extended set of 2^{m+1} signals to the transitions such as to achieve maximum free ED. For codes with up to eight states, this could be accomplished "by hand," and a 16-state code could still be found with the aid of a computer program that checked free ED.

The heuristic design of 8-PSK codes for coded transmission of 2 bit/ T will be discussed in more detail. Uncoded 4-PSK modulation is regarded as a reference system. As shown in Fig. 6, we can view uncoded 4-PSK as coding with a trivial one-state trellis and four "parallel" transitions, to which are assigned from the 8-PSK signal set four signals with largest minimum distance among them, i.e., the signals of subset B0 (or B1). Next consider a two-state trellis. The first code illustrated in Fig. 7 was easily found. Signals of subsets B0 and B1 are assigned to the transitions originating from the first and the second state, respectively, which guarantees that free ED is at least as large as for uncoded 4-PSK modulation. However with two states, it is not possible to have the same property also for transitions joining into one state, and hence the gain in free ED over 4-PSK remains limited to 1.1 dB.

The other 8-PSK codes depicted in Fig. 7 with 4, 8, and 16 states, and gains in free ED of 3 dB, 3.6 dB, and 4.1 dB, respectively, required more effort. Nevertheless after considerable experimentation with various trellis structures and channel-signal assignments, we were convinced that these codes are optimum for the given number of states. The following rules for assigning channel signals were applied:

- 1) all 8-PSK signals should occur with equal frequency

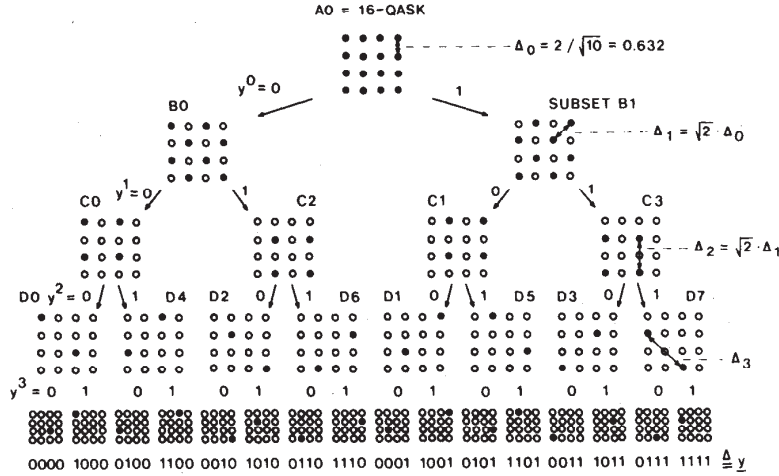


Fig. 5. Partitioning of 16-QASK channel signals into subsets with increasing minimum subset distances ($\Delta_0 < \Delta_1 < \Delta_2 < \Delta_3$; $E\{a_n^2\} = 1$).

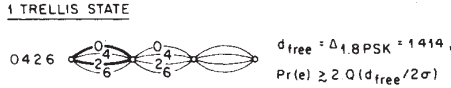


Fig. 6. Uncoded 4-PSK modulation, 2 bit/T.

- and with a fair amount of regularity and symmetry,
- 2) transitions originating from the same state receive signals either from subset B0 or B1,
- 3) transitions joining in the same state receive signals either from subset B0 or B1,
- 4) parallel transitions receive signals either from subset C0 or C1 or C2 or C3.

Rule 1) reflects the intuition that good codes should exhibit regular structures, and rules 2), 3), and 4) guarantee that the ED associated with all single and multiple signal-error events exceeds the free ED of uncoded 4-PSK modulation by at least 3 dB. We have seen that with only two states, rules 2) and 3) cannot be simultaneously satisfied. Since this is not unique to coded 8-PSK, we shall not be further interested in two-state codes.

Note that parallel transitions imply that single signal-error events can occur. This limits achievable free ED to the minimum distance in the subsets of signals assigned to parallel transitions. On the other hand, parallel transitions reduce the "connectivity" in the trellis and thus allow extension of the minimum length of multiple signal-error events. With four states, the trade-off still worked in favor of parallel transitions; the best 4-state 8-PSK code gains 3 dB over 4-PSK, with single signal errors being most likely, whereas codes with distinct transitions to all successor states remained inferior because rules 2) and 3) could not be satisfied simultaneously. With eight and more states, only trellis structures with distinct transitions can be of interest because otherwise free ED gains would remain limited to 3 dB.

The ideas can also be applied to the other modulation forms. As a further example, we consider transmission of 3

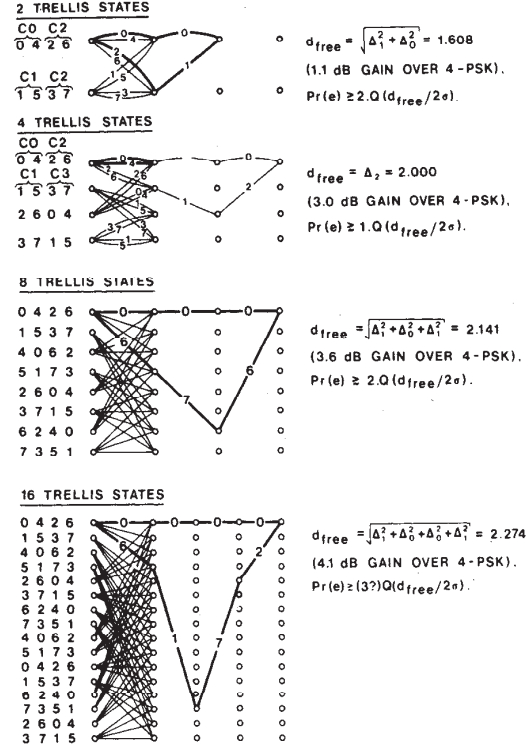


Fig. 7. Coded 8-PSK modulation, 2 bit/T.

bit/T by coded 16-QASK modulation. Uncoded 8-PSK or 8-AMPM modulation is regarded as a reference system. From the preceding discussion and the partitioning of the 16-QASK signal set into subsets shown in Fig. 5, codes follow easily. In the 8-PSK codes of Fig. 7, the 8-PSK signals must only be replaced by the corresponding 16-QASK signal subsets D0-D7 of Fig. 5. An 8-state 16-QASK code obtained in this manner is presented in Fig. 8. The reader must be cautioned, however, about this approach. A similarly obtained 16-state 16-QASK code turned out to

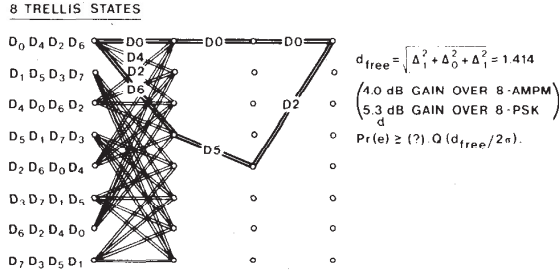


Fig. 8. Coded 16-QASK modulation, 3 bit/T.

have only the same free ED as the 8-state 16-QASK codes.¹

Partitioning of the one-dimensional N -AM signal sets results in minimum subset distances $\Delta_{i+1} = 2 \cdot \Delta_i$, $i = 0, 1, \dots$ (6 dB steps), whereas for two-dimensional N -AMPM and N -QASK signal sets we have $\Delta_{i+1} = \sqrt{2} \cdot \Delta_i$, $i = 0, 1, \dots$ (3 dB steps). Numerical values are given in the code tables of Section V for all signal sets of Fig. 1 and normalized signal power. Partitioning of larger signal sets will soon lead to values of Δ_i that exceed the free ED that one can ever expect to achieve at given code complexity. Therefore it will usually be sufficient—even for very complex codes—to partition a signal set two or three times, and associate the signals of the subsets not further partitioned with parallel transitions in the code trellis.

One further remark is necessary. The specific mapping of coded bits into channel signals indicated in Figs. 4 and 5—which in the case of Fig. 4 leads to a straight binary numbering of the 8-PSK signals—is not important. By permuting subsets, other mappings can be obtained with the same pattern of increasing minimum subset distances. Only this latter property is significant.

IV. MULTILEVEL/PHASE CODES VIEWED AS BINARY CONVOLUTIONAL CODES WITH MAPPING BY SET PARTITIONING

The codes presented in Figs. 7 and 8 have been selected among equivalent codes because with some effort they can be identified as binary convolutional codes of rate $R = m/(m+1)$ generated by feedback-free encoders in cascade with the mapping by set partitioning illustrated by Figs. 4 and 5. Fig. 9 shows the encoders in this form of implementation for constraint lengths $\nu = 2, 3, 4$, corresponding to 4, 8, 16 states.

Using polynomial notation, the binary output sequence $y(D)$ must satisfy the parity check equation:

$$[y^m(D) \cdots y^1(D), y^0(D)] \cdot [H^m(D) \cdots H^1(D), H^0(D)]^T = 0. \quad (8)$$

For the encoders of Fig. 9, the parity check matrices are

$$\begin{aligned} \nu = 2: \quad H(D) &= \begin{bmatrix} (0), & 0, & D, & D^2 + 1 \end{bmatrix}, \\ \nu = 3: \quad H(D) &= \begin{bmatrix} (0), & D, & D^2, & D^3 + 1 \end{bmatrix}, \\ \nu = 4: \quad H(D) &= \begin{bmatrix} D, & D^3 + D^2, & D^4 + D^3 + 1 \end{bmatrix}, \end{aligned} \quad (9)$$

¹Among the codes presented in Section V, there will be a better 16-state 16-QASK code.

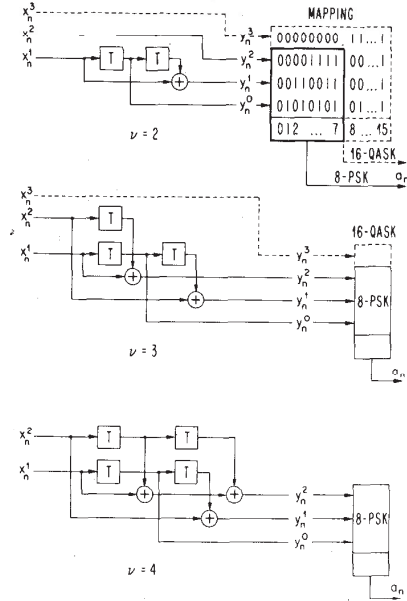


Fig. 9. Realization of 8-PSK and 16-QASK codes by means of minimal convolutional encoders.

where the trivial entry (0) accounts for the additional unchecked bit in coded 16-QASK modulation. Note that the encoders are minimal [13] since $\nu =$ maximum degree of $H^j(D)$, $0 \leq j \leq m$. Considering (9), we observe that the parity-check polynomials have the form ($\nu \geq 2$):

$$H^j(D) = 0, \quad \tilde{m} < j \leq m, \quad (10a)$$

$$H^j(D) = 0 + h_{j-1}^j D^{\nu-1} + \cdots + h_1^j D + 0, \quad 1 \leq j \leq \tilde{m}, \quad (10b)$$

$$H^0(D) = D^\nu + h_{\nu-1}^0 D^{\nu-1} + \cdots + h_1^0 D + 1. \quad (10c)$$

Equation (10a) means that there can be $m - \tilde{m}$ unchecked bits at the binary encoder output. This leads to $2^{m-\tilde{m}}$ parallel transitions between states and hence allows single errors to occur. Note further the difference between (10b) and (10c):

$$h_0^j = h_\nu^j = \begin{cases} 0, & j \neq 0 \\ 1, & j = 0 \end{cases}, \quad (\nu \geq 2). \quad (11)$$

The significance of this condition will be explained below.

Instead of generating code sequences $y(D)$ by minimal encoders, one can also envisage their generation by equivalent systematic encoders with feedback which represent the second canonical form of convolutional encoders [13]. Code sequences $y(D)$ are then generated by

$$[y^m(D) \cdots y^0(D)] = [\dot{x}^m(D) \cdots \dot{x}^1(D)] \cdot \begin{bmatrix} I_m & H^{\tilde{m}}(D)/H^0(D) \\ & \vdots \\ & H^1(D)/H^0(D) \end{bmatrix}, \quad (12)$$

where $\dot{x}(D)$ represents a scrambled version of the input

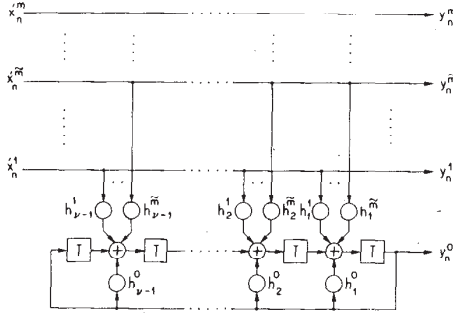


Fig. 10. Systematic convolutional encoder structure with feedback satisfying condition (11).

$x(D)$ to the feedback-free encoders. Because of [11], the rational functions in (12) are realizable. A general realization of (12) with ν delay elements is depicted in Fig. 10, and the specific systematic encoders corresponding to the minimal encoders of Fig. 9 are shown in Fig. 11.

Let $y(D)$ and $y'(D) = y(D) \oplus e(D)$ be two similar binary code sequences where $\oplus$ denotes modulo 2 addition. Since we deal with linear codes, the binary error sequence

$$e(D) = e_k D^k + e_{k+1} D^{k+1} + \dots + e_{k+L} D^{k+L},$$

$$e_k, e_{k+L} \neq 0, \quad L \geq 0 \quad (13)$$

belongs to the set of code sequences. In order to lower-bound the ED between the channel-signal sequences $a(D)$ and $a'(D)$ obtained from $y(D)$ and $y'(D)$, we define the Euclidean weights $w(e_i) = \min d[a(z), a(z \oplus e_i)]$, where minimization goes over all $z = [z^m \dots z^1, z^0]$, and $d[\dots]$ is the ED between the channel signals specified. For the squared ED between $a(D)$ and $a'(D)$, we then obtain

$$\sum_{i=k}^{k+L} d^2[a(y_i), a(y_i \oplus e_i)] \geq \sum_{i=k}^{k+L} w^2(e_i) \triangleq w^2[e(D)]. \quad (14)$$

Theorem: For each $e(D)$ there exists a pair $a(D)$ and $a'(D)$ for which (14) is satisfied with equality.

Proof: Because of symmetry in the channel signal sets $w(e_i) = \min d[a(z), a(z \oplus e_i)]$ is already achieved by letting the last element z^0 in z be arbitrarily 0 or 1, and carrying out the minimization only over $[z^m \dots z^1]$. Since encoding at rate $R = m/(m+1)$ allows any succession of elements $[y_i^m \dots y_i^1]$ to occur (best seen from Fig. 10), there exists for each $e(D)$ a code sequence $y(D)$ that leads in (14) to equality for each individual term of i .

Free ED between channel-signal sequences can therefore be determined in an analogous manner to finding free HD between the binary code sequences $y(D)$. In the appropriate search algorithm that examines all nonzero code sequences $e(D)$ [or $y(D)$], the Hamming weights of e_i must only be replaced by the squared Euclidean weights $w^2(e_i)$. Hence

$$d_{\text{free}}^2 = \min w^2[e(D)] \quad (15a)$$

over all code sequences $e(D) \neq 0$.

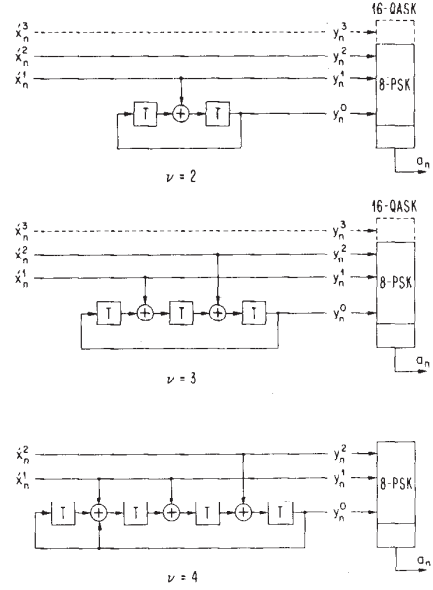


Fig. 11. Equivalent realization of 8-PSK and 16-QASK codes by means of systematic convolutional encoders with feedback.

Let $q(e_i)$ be the number of trailing zeros in e_i , e.g., for $e_i = [e_i^m \dots e_i^1, 1, 0, 0]$ we have $q(e_i) = 2$. From the mapping by set partitioning, one can see that $w(e_i) \geq \Delta_{q(e_i)}$, and that equality holds for almost all e_i . In the 8-PSK mapping illustrated by Fig. 4, we have only found that $w([101]) > \Delta_0$, and that in the 16-QASK mapping in Fig. 5, only $w([1001])$, $w([1101])$, and $w([1111])$ exceed Δ_0 . In all other cases, $w(e_i) = \Delta_{q(e_i)}$ (note $w(0) = \Delta_{q(0)} = 0$). This motivates us to write

$$d_{\text{free}}^2 \gtrsim \Delta_{\text{free}}^2 = \min \sum_{i=k}^{k+L} \Delta_{q(e_i)}^2 \triangleq \min \Delta^2[e(D)] \quad (15b)$$

over all code sequences $e(D) \neq 0$.

The risk taken in setting $d_{\text{free}} = \Delta_{\text{free}}$ is very small because the minimum in (15b) is usually achieved by more than one $e(D)$. This is especially true if $\bar{m} < m$, i.e., higher order bits of e_i are not involved in the parity-check operation. We could never find a code where d_{free} was not equal to Δ_{free} . Therefore we adopt this latter definition of free ED in terms of minimum subset distances, which makes the calculation independent of the exact mapping of coded bits into channel signals as long as the same pattern of these minimum distances results.

To conclude this section, we note another important property of condition (11). From Fig. 10, one can see that multiple signal errors ($L > 0$) must start with $e_k = [e_k^m \dots e_k^1, 0]$ and with $e_{k+L} = [e_{k+L}^m \dots e_{k+L}^1, 0]$. The squared ED associated with these errors is therefore at least $2 \cdot \Delta_1^2$. In other words, the condition guarantees that transitions originating from one state or joining in one state can have signals only from subset B0 or B1 (cf. Section-III, conditions 2) and 3)). We note further that $\bar{m} < m$ implies $d_{\text{free}} \leq \Delta_{\bar{m}}$, since then single-signal errors ($L = 0$) with ED equal to $\Delta_{\bar{m}}$ are possible.

V. SYSTEMATIC SEARCH FOR MORE POWERFUL CODES

The problem to be solved is to find for given constraint length ν and given minimum subset distances $\Delta_0 < \Delta_1 < \dots < \Delta_m$, the systematic encoder with feedback that maximizes $d_{\text{free}} = \Delta_{\text{free}}$. Considering (10), there are $(\tilde{m} + 1) \cdot (\nu - 1)$ coefficients in the parity-check polynomials to be determined. For a feedback-free encoder with a restriction equivalent to (11), one can show that there would be $(\tilde{m} + 1) \cdot (\nu - 1) + \tilde{m}^2 + 1$ generator-polynomial coefficients to decide on. The difference of $\tilde{m}^2 + 1$ accounts for the fact that with the systematic encoder structure, catastrophic encoders are automatically excluded. Another advantage of the systematic encoder structure is that no prior assumptions are necessary on the individual degrees ν_j of the generator polynomials, which for a minimal encoder must satisfy $\sum_{j=1}^m \nu_j = \nu$.

The first step is to determine the value of $\tilde{m}$. From knowledge of d_{free} already achieved by simpler codes, and from the limitation $d_{\text{free}} \leq \Delta_{\tilde{m}}$, the appropriate value follows easily. In fact, we shall see $\tilde{m} = 1$ or 2 only.

The search procedure developed below is basically an exhaustive search with a number of rejection rules. We assume that for good codes (11) is always satisfied. The $\nu - 1$ remaining coefficients in each parity-check polynomial (10b) and (10c) are represented in the form of binary integers

$$0 \leq ih^j = (h_{\nu-1}^j 2^{\nu-2} + h_{\nu-2}^j 2^{\nu-3} + \dots + h_1^j 2^0) \leq 2^{\nu-1} - 1, \quad 0 \leq j \leq \tilde{m}. \quad (16)$$

Definition: An incomplete code at level l , $1 \leq l < \tilde{m}$ is specified by $ih^l = [ih^l, ih^{l-1}, \dots, ih^0]$. We let this be equivalent to setting encoder inputs $x^{l+1}(D), \dots, x^{\tilde{m}}(D)$ to zero for a complete code specified by $ih^{\tilde{m}}$.

In the search program, we increment in the outermost loop ih^0 from 0 to $2^{\nu-1} - 1$, in the next inner loop ih^1 from 0 to $2^{\nu-1} - 1$, etc. Whenever at level l a code is rejected by one of the following rules, we skip inner loops at levels $l + 1$ to $\tilde{m}$.

Rule 1: Reversing time does not change the distance properties of a code. Let ih_{rev}^j be the bit-reversed binary integer ih^j , e.g., $(0111)_2 \leftrightarrow (1110)_2$. It follows from the order in which the ih^j are varied in the search program that a code can be rejected at level l if $ih_{\text{rev}}^l < ih^l$ and $ih_{\text{rev}}^j = ih^j$ for $0 \leq j < l$.

Rule 2: Let $H'(D)$ be derived from $H(D)$ by

$$H'(D) = [H^m(D), \dots, H^t(D), \dots, H^t(D) \oplus H^s(D), \dots, H^0(D)], \quad 0 \leq s < t \leq m. \quad (17)$$

If

$$e(D) = [e^m(D), \dots, e^t(D), \dots, e^s(D), \dots, e^0(D)] \quad (18)$$

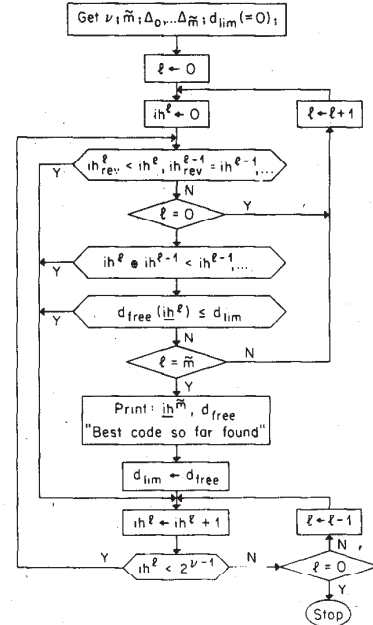


Fig. 12. Block diagram of code search program.

satisfies $e(D) \cdot H^{\text{tr}}(D) = 0$, then

$$e'(D) = [e^m(D), \dots, e^t(D) \oplus e^s(D), \dots, e^s(D), \dots, e^0(D)] \quad (19)$$

satisfies $e'(D) \cdot H^{\text{tr}}(D) = 0$. Since the element vectors e_l and e'_l of (18) and (19) exhibit the same number of trailing zeros, the same lower bound on ED expressed in subset distances $\Delta_{q(e_i)}$ applies. We conclude that the codes defined by $H(D)$ and $H'(D)$ have identical free ED. Again taking into account the order in which the ih^j are varied in the search program, a code can be rejected at level l if one of the following conditions is satisfied:

$$\begin{aligned} l \geq 1: & \quad ih^l \oplus ih^{l-1} < ih^{l-1}, \\ l \geq 2: & \quad ih^l \oplus ih^{l-2} < ih^{l-2}, \\ & \quad ih^l \oplus ih^{l-1} \oplus ih^{l-2} < ih^{l-2}, \\ l \geq 3: & \quad ih^l \oplus ih^{l-3} < ih^{l-3}, \\ & \quad ih^l \oplus ih^{l-2} \oplus ih^{l-3} < ih^{l-3}, \\ & \quad ih^l \oplus ih^{l-1} \oplus ih^{l-2} < ih^{l-2}, \\ & \quad ih^l \oplus ih^{l-1} \oplus ih^{l-2} \oplus ih^{l-3} < ih^{l-3}, \text{ etc.} \end{aligned} \quad (20)$$

Here, $ih^t \oplus ih^s$ denotes the integer obtained by modulo 2 addition of ih^t and ih^s bit by bit.

Rule 3: A code is rejected at level l if its free ED, $d_{\text{free}}^{(l)}$, does not exceed the largest value $d_{\text{free}}^{(\tilde{m})} = d_{\text{lim}}$ found among previously inspected complete codes. Initially, d_{lim} can be set to zero, or in order to save time, to the free ED of the best code previously found with smaller constraint length ν .

TABLE I
 CODES FOR AMPLITUDE MODULATION

ν	$\tilde{m}$	$H^0(D)$	$H^1(D)$	$H^2(D)$	$d_{\text{free}}^2/\Delta_1^2$	$m=1$ $G_{4AM/2AM}$	$m=2$ $G_{8AM/4AM}$	$m=3$ $G_{16AM/8AM}$
2	1	5 ₈	2 ₈	-	2.25	~2.5 dB	~3.3 dB	~3.5 dB
3	1	13	04	-	2.50	3.0	3.8	3.9
4	1	23	04	-	2.75	3.4	4.2	4.3
5	1	45	10	-	3.25	4.2	5.0	5.1
6	1	103	024	-	3.50	4.5	5.2	5.4
7	1	235	126	-	4.00	5.1	5.8	6.0
8	1	515	262	-	4.25	5.3	-	-
9	1	1017	0342	-	4.50	5.6	-	-
* 10	1	2051	1536	-	4.75	5.8	-	-
* 11	1	4017	1602	-	5.00	6.0	-	-
Δ 8	2	403	170	000	4.00	-	5.8	6.0
Δ 9	2	1003	0134	0000	4.00	-	5.8	6.0
* Δ 10	2	4003	0134	0000	4.00	-	5.8	6.0

$E\{ a_n^2 \} = 1:$	$\Delta_1^2(4AM)/\Delta_0^2(2AM) = 4/5$ (-0.97 dB)
$\Delta_0(N-AM) = [12/(N^2-1)]^{1/2}$	$\Delta_1^2(8AM)/\Delta_0^2(4AM) = 20/21$ (-0.21 dB)
$\Delta_1(N-AM) = 2\Delta_{i-1}(N-AM), i=1, 2, \dots$	$\Delta_1^2(16AM)/\Delta_0^2(8AM) = 84/85$ (-0.05 dB)

*Search not completed.
 Δ No improvement obtained.

Further rules are conceivable; for example, a code is rejected if

$$\text{GCD}\{H^{\tilde{m}}(D), \dots, H^0(D)\} \neq 1. \quad (21)$$

This latter rule was not implemented.

A block diagram of the code-search program is presented in Fig. 12. For the calculation of $d_{\text{free}}^{(l)}$, the bidirectional search algorithm in the form discussed by Larsen [14] was adopted. For computing in this algorithm distance increments of state extensions, the systematic encoder structure offers the same computational advantages which Paaske [12] obtained by regarding the syndrome former states instead of the ordinary states of feedback-free encoders.

The codes obtained by the program are listed in Tables I–III. For each constraint length ν , a limit of 30 min CPU time was set, and no particular effort for program optimization was made (IBM/370–158, PL/I program without special assembler written routines). Parity check polynomials are presented in the Tables in octal form, e.g., $H^0(D) = D^6 + D^2 + 1 \triangleq (001\,000\,101)_2 \triangleq (105)_8$. Free ED is first given in the normalized form $d_{\text{free}}^2/\Delta_1^2(\text{coded})$. Note that for normalized signal power $E\{|a_n^2|\} = 1$, we have $\Delta_1(\text{coded}) \simeq \Delta_0(\text{uncoded})$. The asymptotic coding gain for each specific coded/uncoded comparison is then computed by replacing $\Delta_1(\text{coded})$ by the appropriate $\Delta_0(\text{uncoded})$

 TABLE II
 CODES FOR 8-PSK MODULATION

ν	$\tilde{m}$	$H^0(D)$	$H^1(D)$	$H^2(D)$	$d_{\text{free}}^2/\Delta_1^2$	$m=2$ $G_{8PSK/4PSK}$
2	1	5 ₈	2 ₈	-	2.000	~ 3.0 dB
3	2	11	02	04	2.293	3.6
4	2	23	04	16	2.586	4.1
5	2	45	16	34	2.879	4.6
6	2	105	036	074	3.000	4.8
7	2	203	014	016	3.172	5.0
8	2	405	250	176	3.465	5.4
9	2	1007	0164	0260	3.758	5.7
* Δ 10	2	2003	0164	0770	3.758	5.7

$E\{ a_n^2 \} = 1:$	$\Delta_1^2(8PSK)/\Delta_0^2(4PSK)$
$\Delta_0(4PSK) = \sqrt{2}, \Delta_0(8PSK) = 2\sin(\pi/8),$	$= 1$ (0 dB)
$\Delta_1(8PSK) = \sqrt{2}, \Delta_2(8PSK) = 2$	

*Search not completed.
 Δ No improvement obtained.

and expressing the above ratio in decibels. For example,

$$G_{N-AM/(N/2)-AM} = 10 \cdot \log_{10} \left\{ d_{\text{free}}^2(N-AM)/\Delta_0^2[(N/2)-AM] \right\}. \quad (22)$$

The tables show that increasing ν does not always lead to a code with larger d_{free} , especially if coding gains are already in the order of 6 dB. The code search could have been extended to larger values of ν by optimizing critical program sections, allowing longer program runs, and perhaps by including additional rejection rules. In view of the exponential increase in code complexity and the small additional gains in free ED to be expected, we were content with the codes already found.

VI. SIMULATION RESULTS

Coding gains have so far been expressed in terms of larger free ED, which is adequate for high SNR. In this section, we show simulation results obtained at moderate SNR. We concentrate mainly on error-event probability which is relevant for most blockwise data communication. This makes the result independent of the specific mapping of information bit sequences into channel signal sequences. In a practical system, several factors can influence this mapping: the particular bit labeling of branches in signal-set partitioning, the realization of the encoder in the form of a feedback-free minimal encoder or a systematic encoder with feedback, phase-differential coding in order to resolve phase ambiguity, and scrambling. Results on bit-error

TABLE III
CODES FOR AMPLITUDE/PHASE MODULATION

ν	$\tilde{m}$	$H^0(D)$	$H^1(D)$	$H^2(D)$	$d_{\text{free}}^2/\Delta_1^2$	$G^m = 2$ 8AMP/4PSK	$G^m = 3$ 16QASK/ 8AMP(8PSK)	$G^m = 4$ 32AMP/ 16QASK	$G^m = 5$ 64QASK/ 32AMP
2	1	5 ₈	2 ₈	-	2.0	~2.0 dB	~3.0(4.4) dB	~2.8 dB	~3.0 dB
3	2	11	02	04	2.5	3.0	4.0(5.3)	3.8	4.0
4	2	23	04	16	3.0	3.8	4.8(6.1)	4.6	4.8
Δ 5	2	41	06	10	3.0	3.8	4.8(6.1)	4.6	4.8
6	2	101	016	064	3.5	4.5	5.4(6.8)	5.2	5.4
7	2	203	014	042	4.0	5.1	6.0(7.4)	5.8	6.0
Δ 8	2	401	056	354	4.0	5.1	6.0(7.4)	5.8	6.0
9	2	1001	0346	0510	4.5	5.6	-	-	-
*10	2	2003	0164	0770	5.0	6.0	-	-	-

$E\{ a_n^2 \} = 1:$	$\Delta_1^2(8\text{AMP})/\Delta_0^2(4\text{PSK}) = 4/5$	(- 0.97 dB)
$\Delta_0(8\text{AMP}) = 2/\sqrt{5}, \Delta_0(16\text{QASK}) = \sqrt{2}/5,$	$\Delta_1^2(16\text{QASK})/\Delta_0^2(8\text{AMP}) = 1$	(0.0 dB)
$\Delta_0(32\text{AMP}) = 2/\sqrt{21}, \Delta_0(64\text{QASK}) = \sqrt{2}/21,$	$\Delta_1^2(16\text{QASK})/\Delta_1^2(8\text{PSK}) = 1/(5\sin^2(\pi/8))$	(+ 1.35 dB)
$\Delta_i(x) = \sqrt{2} \cdot \Delta_{i-1}(x), i = 1, 2, \dots$	$\Delta_1^2(32\text{AMP})/\Delta_0^2(16\text{QASK}) = 20/21$	(- 0.21 dB)
	$\Delta_1^2(64\text{QASK})/\Delta_0^2(32\text{AMP}) = 1$	(0.0 dB)

*Search not completed.

Δ No improvement obtained.

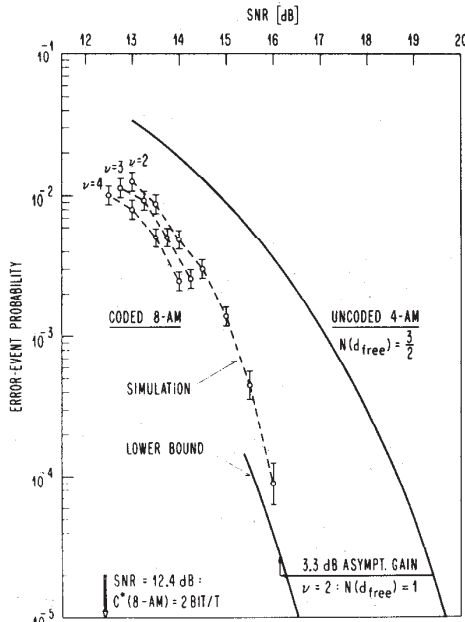


Fig. 13. Error-event performance of coded 8-AM versus uncoded 4-AM, 2 bit/T.

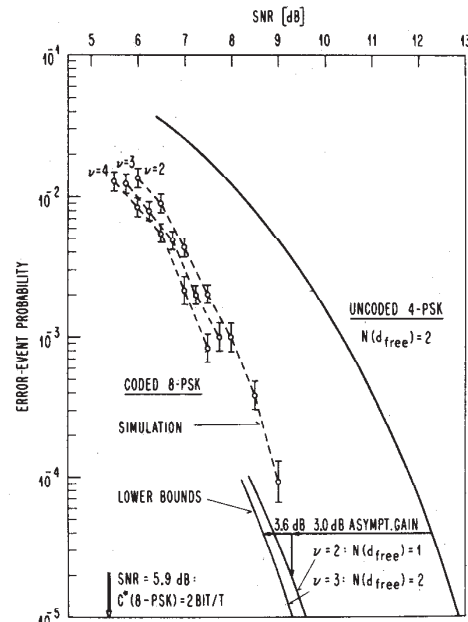


Fig. 14. Error-event performance of coded 8-PSK versus uncoded 4-PSK, 2 bit/T.

probability are therefore given only for one particular example.

In the simulation programs, the length of the path histories in the Viterbi decoding algorithms was $M = 6\nu$ (decision delay). An error event was counted when a false channel signal was decoded following a state that still belonged to the correct path through the code trellis. Figs. 13–15 show error-event frequency versus SNR for coded 8-AM, 8-PSK, and 16-QASK in comparison with computed error-event probability of 4-AM, 4-PSK, and 8-PSK

or 8-AMPM, respectively. The simulation results are given with 90 percent confidence intervals. Systematic encoders with the parity-check polynomials given in Tables I–III for $\nu = 2, 3$, and 4 were used. For the simpler codes, where $N(d_{\text{free}})$ could still be determined by inspection of the trellis diagrams, lower bounds on error-event probability are included. Also indicated in Figs. 13–15 is the SNR at which channel capacity of the expanded signal sets (cf. Figs. 2(a) and (b)) equals the transmission rate in bit/T. This illustrates that by simple 4-state codes a significant

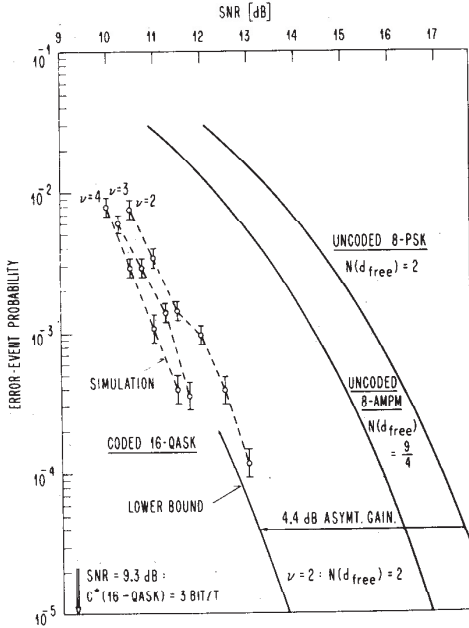
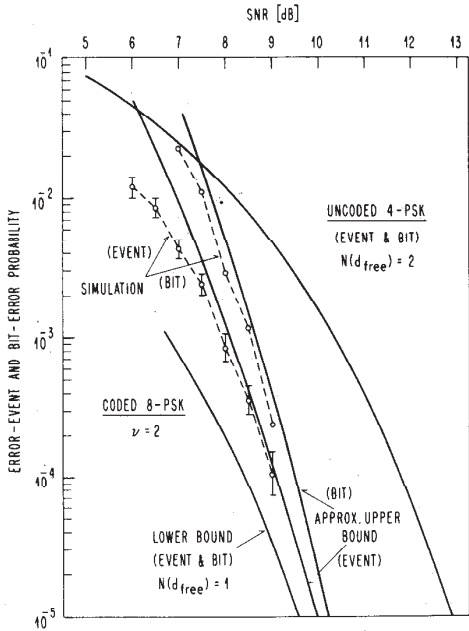
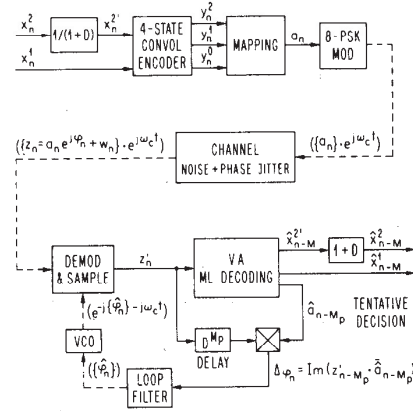


Fig. 15. Error-event performance of coded 16-QASK versus uncoded 8-PSK and 8-APPM, 3 bit/T.


 Fig. 16. Error-event and bit-error performance of coded 8-PSK ($\nu = 2$, minimal encoder) and uncoded 4-PSK, 2 bit/T.

portion of the totally possible improvement of performance is already achieved, and that further improvements by using more complex codes ($\nu = 3, 4, \dots$) tend to be relatively expensive.

Fig. 16 shows error-event and bit-error frequency of coded 8-PSK using the 4-state minimal encoder depicted in Fig. 9. The simplicity of this code permitted the calculation of lower and appropriate upper bounds on error-event and


 Fig. 17. Coded 8-PSK transmission system with carrier-phase tracking loop and differential coding/decoding to resolve 180° phase ambiguity.

bit-error probabilities (upper-bounds will not be discussed). It can be observed that the two error probabilities become identical for high SNR because in this particular case, the free ED occurs for single-signal error events with only one wrong bit per event, and because no differential coding and scrambling is used.

VII. CARRIER-PHASE TRACKING

In the preceding sections, ideal timing and carrier-phase tracking were assumed to be available in the receiver. Whereas timing synchronization can be accomplished in the usual manner by signal squaring and filtering with a narrow bandpass filter tuned to $1/T$ Hz, carrier-phase synchronization can pose a somewhat harder problem. In this section, we restrict attention to coded 8-PSK modulation using the 4-state encoder shown in Fig. 9, and measure the phase tracking performance of a simulated receiver. The problems encountered may be regarded as typical of other cases of coded carrier modulation. Instead of (1), we assume that the receiver deals with

$$z_n = a_n \cdot e^{j\varphi_n} + w_n \quad (23)$$

where φ_n is a slowly varying carrier phase *a priori* unknown in the receiver. Fig. 17 illustrates the tracking of φ_n by a decision-directed carrier-phase loop that uses tentative decisions of the Viterbi decoding algorithm: $\hat{\varphi}_n$ represents an estimate of φ_n , and $\hat{a}_{n-M_p}$ denotes a tentative signal decision that is obtained by backtracking in the decoding algorithm the survivor path from the most likely survivor state at time n , $M_p \leq M$ times. The input to the decoding algorithm becomes

$$z'_n = a_n \cdot e^{j(\varphi_n - \hat{\varphi}_n)} + w'_n. \quad (24)$$

A measured phase error

$$\Delta\varphi_n = \angle(z'_{n-M_p} \cdot \hat{a}_{n-M_p}) \simeq \text{Im}(z'_{n-M_p} \cdot \bar{\hat{a}}_{n-M_p}), \quad (|a_n| = 1) \quad (25)$$

is used to predict φ_{n+1} by a first-order filter

$$\hat{\varphi}_{n+1} = \hat{\varphi}_n + \gamma \cdot \Delta\varphi_n, \quad 0 < \gamma < 1. \quad (26)$$

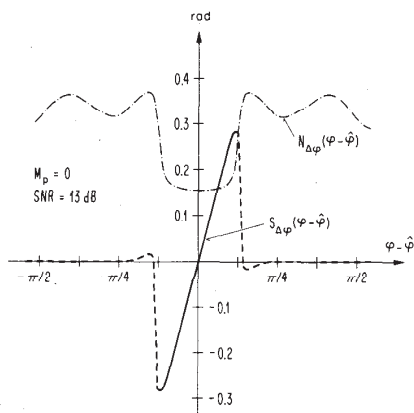


Fig. 18. Carrier-phase tracking loop characteristics: S-curve and standard deviation of loop noise in the case of 4-state coded 8-PSK modulation.

Several questions arise:

- 1) How should the tentative decision delay M_p be chosen? Larger M_p lead to more reliable decisions, but also increase loop delay.
- 2) Which value should be used for the loop gain γ ?
- 3) How can phase ambiguity be dealt with?

In simulating the arrangement of Fig. 17, $\{\varphi_n\}$ has been modeled as a process with independent Gaussian increments. For M_p we found that shorter loop delay is preferable to having more reliable tentative decisions, hence $M_p = 0$ is a good choice. Further insight in the tracking behavior was obtained by measuring the mean value $S_{\Delta\varphi} = E(\Delta\varphi_n)$ and the standard deviation $N_{\Delta\varphi} = |E(\Delta\varphi_n - S_{\Delta\varphi})^2|^{1/2}$ as a function of $\varphi_n - \hat{\varphi}_n$, by setting in the simulation program $\hat{\varphi}_n = \gamma = 0$ and varying φ_n . The so-called S-curve of the loop, $S_{\Delta\varphi}$, and $N_{\Delta\varphi}$ are presented in Fig. 18; these indicate normal tracking behavior within the range $|\varphi_n - \hat{\varphi}_n| < \pi/8$. Outside this range, the loop performs a random walk driven by loop noise. With an appropriate value of γ , locking into the normal tracking range was consistently observed after some random delay.

The curves depicted in Fig. 18 are periodic with period π , which results in a 180° phase ambiguity. In this particular case, phase ambiguity can be resolved by differential-encoding/decoding of the most significant bit as indicated in Fig. 17. Different approaches may be required for other codes.

VIII. CONCLUSION

By computing the channel capacity of multilevel/phase modulation schemes, we found that doubling the number of channel-signals makes most of the theoretically possible improvement accessible. From simple hand-designed codes, we deduced a general structure of multilevel/phase encoders as binary convolutional encoders of rate $R = m/(m+1)$ followed by mapping of coded bits into channel

signals by "set partitioning." This mapping allowed the derivation of an efficient lower-bound on Euclidean distance between resulting channel-signal sequences. The bound was applied in a search program for codes that achieve maximum free Euclidean distance for given constraint length of the convolutional encoder.

The general finding of this paper is that compared with uncoded modulation, the same amount of information can be transmitted within the same bandwidth with coding gains of 3–4 dB by simple hand-designed codes with four to eight states. Improvements in the order of 6 dB require codes with about 2^{10} states. No significant difference was found between the performance gains achievable with one- and two-dimensional coded modulation. It appears unlikely therefore that using signal sets defined in more than two dimensions could have advantages. We have only studied $R = m/(m+1)$ codes. It could be that by expanding channel-signal sets more than twice, codes could be found with better free ED for given constraint length than the codes presented in this paper, although little increase in channel capacity would occur. We leave this as an open problem for further investigations.

Coding usually leads to an additional synchronization problem. We have demonstrated how to perform carrier-phase tracking for one particular example.

ACKNOWLEDGMENT

The author wishes to thank J. L. Massey for his continuous interest in this work and many helpful suggestions on presenting the material.

REFERENCES

- [1] J. L. Massey, "Coding and modulation in digital communications," *Proc. 1974 Int. Zurich Seminar on Digital Comm.*, Zurich, Switzerland, pp. E2(1)–(4), Mar. 1974.
- [2] G. D. Forney, Jr., "The Viterbi algorithm," *Proc. IEEE*, vol. 61, pp. 268–278, Mar. 1973.
- [3] G. Ungerboeck and I. Csajka, "On improving data-link performance by increasing the channel alphabet and introducing sequence coding," *1976 Int. Symp. Inform. Theory*, Ronneby, Sweden, June 1976.
- [4] J. P. Odenwalder, A. J. Viterbi, I. M. Jacobs, and J. A. Heller, "Study on information transfer optimization for communication satellites," Final Rep. NASA CR No. 114561, Linkabit Corporation, Section 3.3, Jan. 1973.
- [5] A. Digeon, "On improving bit error probability of QPSK and 4-level amplitude modulation systems by convolutional coding," *IEEE Trans. Commun.*, vol. COM-25, pp. 1238–1239, Oct. 1977.
- [6] J. B. Anderson and R. de Buda, "Better phase-modulation performance using trellis phase codes," *Electron. Lett.*, vol. 12, pp. 587–588, Oct. 1976.
- [7] J. B. Anderson and D. P. Taylor, "A bandwidth-efficient class of signal-space codes," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 703–712, Nov. 1978.
- [8] T. Aulin, N. Rydbeck, and C. E. Sundberg, "Transmitter and receiver structures for M-ary partial response," in *Proc. 1980 Int. Zurich Seminar on Digital Comm.*, Zurich, Switzerland, pp. A2–1–6, Mar. 1980.
- [9] D. P. Taylor and H. C. Chan, "A simulation study of some bandwidth-efficient modulation techniques," *Comm. Research Lab.*, McMaster University, Canada, Rep. No. CRL-68, Nov. 1979.

- [10] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. New York, Wiley, 1965, pp. 318–319.
- [11] R. G. Gallager, *Information Theory and Reliable Communication*. New York, Wiley, 1968, p. 74.
- [12] E. Paaske, "Short binary convolutional codes with maximal free distance for rates $2/3$ and $3/4$," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 683–689, Sept. 1974.
- [13] G. D. Forney, Jr., "Convolutional codes I: algebraic structure," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 720–738, Nov. 1970.
- [14] K. J. Larsen, "Comments on 'An efficient algorithm for computing free distance'," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 437–439, May 1972.

Finite Sampling Approximations for Non-Band-Limited Signals

STAMATIS CAMBANIS MEMBER, IEEE, AND MUHAMMAD K. HABIB

Abstract—Finite sampling approximations, along with bounds on the approximation error, are derived for certain deterministic and random signals which are not band-limited.

I. STATEMENT AND DISCUSSION OF RESULTS

THIS PAPER derives finite sampling approximations and their rates of convergence for deterministic and random signals which are not band-limited. The merit of these results lies in the fact that in many practical situations the signals under consideration are not band-limited, and only a finite number of samples are available. We thus study conditions under which a finite sum of the form

$$\sum_{n=-N(W)}^{N(W)} f\left(\frac{n}{2W}\right) \frac{\sin \pi(2Wt - n)}{\pi(2Wt - n)} \quad (1)$$

converges to the non-band-limited function $f(t)$ as the sampling rate W tends to infinity, and as the number $2N(W) + 1$ of samples used, which depends on the sampling rate W , also tends to infinity with W . We also determine the speed of the convergence.

The problem of approximating a non-band-limited signal by an infinite series of the form (1), i.e., $N(W) = \infty$, has been considered for deterministic signals in [2], for random stationary signals in [5], for time-limited deterministic signals in [6], and for certain analytic signals in [21] and [14].

Manuscript received April 22, 1980; revised January 13, 1981. This work was supported in part by the Air Force Office of Scientific Research under Grant AFOSR-80-0080 and in part by a scholarship of the Arab Republic of Egypt. This paper was presented at the 14th Annual Conference on Information Sciences and Systems, Princeton, NJ, March 26–28, 1980.

S. Cambanis is with the Department of Statistics, University of North Carolina, Chapel Hill, NC 27514.

M. K. Habib is with the Department of Biostatistics, School of Public Health, University of North Carolina, Chapel Hill, NC 27514.

There is, of course, extensive literature on bounds for the truncation error in the approximation of a band-limited signal by a finite sum of the form (1), in chronological order [22], [10], [18], [3], [4], [17], [1]. In this case, W is fixed, and N is independent of W and tends to infinity. For a review of the sampling theorem, see [11].

Here we consider the approximation of non-band-limited signals by finite sums of the form (1). This problem has been considered for certain analytic deterministic signals in [21], [14]; for certain deterministic as well as random signals which have at least two derivatives in [13]; for certain deterministic signals whose derivatives satisfy a Lipschitz condition in [7]; for certain deterministic signals with multidimensional parameter which have mixed partial derivatives in [19]; and for deterministic signals which are integrable, have integrable Fourier transform, and fall off faster than $\text{const.}|t|^{-\alpha}$, $\alpha > 2$, in [9]. The deterministic and random signals considered in this paper satisfy much less stringent conditions. Although we consider only signals with a one-dimensional parameter, generalization to multi-dimensional parameter signals should be feasible.

The results for deterministic signals are stated in Theorem 1 and Theorem 2 and its corollary. The results for random signals are stated in Theorems 3–5. The derivations of the results are given in Section II.

We begin by considering a (Cezàro) version of (1), given by (4) (see also [16]). The (Cezàro) coefficients ensure that (4) converges to $f(t)$, provided a sufficient number of samples is used (Theorem 1). In contrast to this general and simple result, the convergence of (1) to $f(t)$ is a more complicated matter requiring additional assumptions on the signal and the number of samples required. The difference between (1) and (4) is that while (4) leads to the nonnegative Fejér kernel, (1) leads to the more intractable Dirichlet kernel (cf. (19), (20) and (28), (29)).