Trellis-Coded Modulation with Redundant Signal Sets Part I: Introduction

Gottfried Ungerboeck

Simple four-state trellis-coded modulation (TCM) schemes improve the robustness of digital transmission against additive noise by 3 dB without reducing data rate or requiring more bandwidth than conventional uncoded modulation schemes. With more complex schemes, coding gains up to 6 dB can be achieved. This article describes how TCM works

0163-6804/87/0002-0005 $01.00 © 1987 IEEE

Trellis-Coded Modulation (TCM) has evolved over the past decade as a combined coding and modulation technique for digital transmission over band-limited channels. Its main attraction comes from the fact that it allows the achievement of significant coding gains over conventional uncoded multilevel modulation without compromising bandwidth efficiency. The first TCM schemes were proposed in 1976 [1]. Following a more detailed publication [2] in 1982, an explosion of research and actual implementations of TCM took place, to the point where today there is a good understanding of the theory and capabilities of TCM methods. In Part I of this two-part article, an introduction into TCM is given. The reasons for the development of TCM are reviewed, and examples of simple TCM schemes are discussed. Part II [15] provides further insight into code design and performance, and addresses recent advances in TCM.

TCM schemes employ redundant nonbinary modulation in combination with a finite-state encoder which governs the selection of modulation signals to generate coded signal sequences. In the receiver, the noisy signals are decoded by a soft-decision maximum-likelihood sequence decoder. Simple four-state TCM schemes can improve the robustness of digital transmission against additive noise by 3 dB, compared to conventional uncoded modulation. With more complex TCM schemes, the coding gain can reach 6 dB or more. These gains are obtained without bandwidth expansion or reduction of the effective information rate as required by traditional error-correction schemes. Shannon's information theory predicted the existence of coded modulation schemes with these characteristics more than three decades ago. The development of effective TCM techniques and today's signal-processing technology now allow these gains to be obtained in practice.

Signal waveforms representing information sequences are most impervious to noise-induced detection errors if they are very different from each other. Mathematically, this translates into the requirement that signal sequences should have large distance in Euclidean signal space. The essential new concept of TCM that led to the aforementioned gains was to use signal-set expansion to provide redundancy for coding, and to design coding and signal-mapping functions jointly so as to maximize directly the "free distance" (minimum Euclidean distance) between coded signal sequences. This allowed the construction of modulation codes whose free distance significantly exceeded the minimum distance between uncoded modulation signals, at the same information rate, bandwidth, and signal power. The term "trellis" is used because these schemes can be described by a state-transition (trellis) diagram similar to the trellis diagrams of binary convolutional codes. The difference is that in TCM schemes, the trellis branches are labeled with redundant nonbinary modulation signals rather than with binary code symbols.

The basic principles of TCM were published in 1982 [2]. Further descriptions followed in 1984 [3-6], and coincided with a rapid transition of TCM from the research stage to practical use. In 1984, a TCM scheme with a coding gain of 4 dB was adopted by the International Telegraph and Telephone Consultative Commit-

tee (CCITT) for use in new high-speed voiceband modems [5,7,8]. Prior to TCM, uncoded transmission at 9.6 kbit/s over voiceband channels was often considered as a practical limit for data modems. Since 1984, data modems have appeared on the market which employ TCM along with other improvements in equalization, synchronization, and so forth, to transmit data reliably over voiceband channels at rates of 14.4 kbit/s and higher. Similar advances are being achieved in transmission over other bandwidth-constrained channels. The common use of TCM techniques in such applications, as satellite [9-11], terrestrial microwave, and mobile communications, in order to increase throughput rate or to permit satisfactory operation at lower signal-to-noise ratios, can be safely predicted for the near future.

Classical Error-Correction Coding

In classical digital communication systems, the functions of modulation and error-correction coding are separated. Modulators and demodulators convert an analog waveform channel into a discrete channel, whereas encoders and decoders correct errors that occur on the discrete channel.

In conventional multilevel (amplitude and/or phase) modulation systems, during each modulation interval the modulator maps m binary symbols (bits) into one of $M = 2^m$ possible transmit signals, and the demodulator recovers the m bits by making an independent M-ary nearest-neighbor decision on each signal received. Figure 1 depicts constellations of real- or complex-valued modulation amplitudes, henceforth called signal sets, which are commonly employed for one- or two-dimensional M-ary linear modulation. Two-dimensional carrier modulation requires a bandwidth of 1/T Hz around the carrier frequency to transmit signals at a modulation rate of 1/T signals/sec (baud) without intersymbol interference. Hence, two-dimensional 2^m-ary modulation systems can achieve a spectral efficiency of about m bit/sec/Hz. (The same spectral efficiency is obtained with one-dimensional $2^{m\,2}$-ary baseband modulation.)

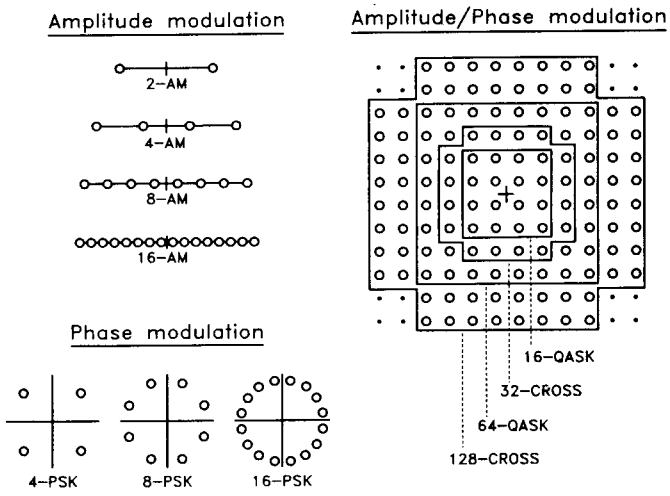

Fig. 1. Signal sets for one-dimensional amplitude modulation, and two-dimensional phase and amplitude/phase modulation.

Conventional encoders and decoders for error correction operate on binary, or more generally Q-ary, code symbols transmitted over a discrete channel. With a code of rate $k/n < 1$, $n - k$ redundant check symbols are appended to every k information symbols. Since the decoder receives only discrete code symbols, Hamming distance (the number of symbols in which two code sequences or blocks differ, regardless of how these symbols differ) is the appropriate measure of distance for decoding and hence for code design. A minimum Hamming distance d^H_{min}, also called "free Hamming distance" in the case of convolutional codes, guarantees that the decoder can correct at least $[(d^H_{min} - 1)/2]$ code-symbol errors. If low signal-to-noise ratios or non-stationary signal disturbance limit the performance of the modulation system, the ability to correct errors can justify the rate loss caused by sending redundant check symbols. Similarly, long delays in error-recovery procedures can be a good reason for trading transmission rate for forward error-correction capability.

Generally, there exist two possibilities to compensate for the rate loss: increasing the modulation rate if the channel permits bandwidth expansion, or enlarging the signal set of the modulation system if the channel is band-limited. The latter necessarily leads to the use of nonbinary modulation ($M > 2$). However, when modulation and error-correction coding are performed in the classical independent manner, disappointing results are obtained.

As an illustration, consider four-phase modulation (4-PSK) without coding, and eight-phase modulation (8-PSK) used with a binary error-correction code of rate 2/3. Both systems transmit two information bits per modulation interval (2 bit/sec/Hz). If the 4-PSK system operates at an error rate of 10^{-5}, at the same signal-to-noise ratio the "raw" error rate at the 8-PSK demodulator exceeds 10^{-2} because of the smaller spacing between the 8-PSK signals. Patterns of at least three bit errors must be corrected to reduce the error rate to that of the uncoded 4-PSK system. A rate-2/3 binary convolutional code with constraint length $\nu = 6$ has the required value of $d^H_{min} = 7$ [12]. For decoding, a fairly complex 64-state binary Viterbi decoder is needed. However, after all this effort, error performance only breaks even with that of uncoded 4-PSK.

Two problems contribute to this unsatisfactory situation.

Soft-Decision Decoding and Motivation for New Code Design

One problem in the coded 8-PSK system just described arises from the independent "hard" signal decisions made prior to decoding which cause an irreversible loss of information in the receiver. The remedy for this problem is soft-decision decoding, which means that the decoder operates directly on unquantized "soft" output samples of the channel. Let the samples be $r_n = a_n + w_n$ (real- or complex-valued, for one- or two-dimensional modulation, respectively), where the a_n are the discrete signals sent by the modulator, and the w_n represent samples of an additive white Gaussian noise process. The decision rule of the optimum sequence decoder is to

determine, among the set C of all coded signal sequences which a cascaded encoder and modulator can produce, the sequence $\{\hat{a}_n\}$ with minimum squared Euclidean distance (sum of squared errors) from $\{r_n\}$, that is, the sequence $\{\hat{a}_n\}$ which satisfies

$$|r_n - \hat{a}_n|^2 = \underset{\{\hat{a}_n\}\epsilon C}{Min} \sum |r_n - a_n|^2.$$

The Viterbi algorithm, originally proposed in 1967 [13] as an "asymptotically optimum" decoding technique for convolutional codes, can be used to determine the coded signal sequence $\{\hat{a}_n\}$ closest to the received unquantized signal sequence $\{r_n\}$ [12,14], provided that the generation of coded signal sequences $\{a_n\}\epsilon C$ follows the rules of a finite-state machine. However, the notion of "error-correction" is then no longer appropriate, since there are no hard-demodulator decisions to be corrected. The decoder determines the most likely coded signal sequence directly from the unquantized channel outputs.

The most probable errors made by the optimum soft-decision decoder occur between signals or signal sequences $\{a_n\}$ and $\{b_n\}$, one transmitted and the other decoded, that are closest together in terms of squared Euclidean distance. The minimum squared such distance is called the squared "free distance:"

$$d^2_{free} = \underset{\{a_n\}\neq\{b_n\}}{Min} \sum |a_n - b_n| \; ; \quad \{a_n\}, \{b_n\} \epsilon C.$$

When optimum sequence decisions are made directly in terms of Euclidean distance, a second problem becomes apparent. Mapping of code symbols of a code optimized for Hamming distance into nonbinary modulation signals does not guarantee that a good Euclidean distance structure is obtained. In fact, generally one cannot even find a monotonic relationship between Hamming and Euclidean distances, no matter how code symbols are mapped.

For a long time, this has been the main reason for the lack of good codes for multilevel modulation. Squared Euclidean and Hamming distances are equivalent only in the case of binary modulation or four-phase modulation, which merely corresponds to two orthogonal binary modulations of a carrier. In contrast to coded multilevel systems, binary modulation systems with codes optimized for Hamming distance and soft-decision decoding have been well established since the late 1960s for power-efficient transmission at spectral efficiencies of less than 2 bit/sec/Hz.

The motivation of this author for developing TCM initially came from work on multilevel systems that employ the Viterbi algorithm to improve signal detection in the presence of intersymbol interference. This work provided him with ample evidence of the importance of Euclidean distance between signal sequences. Since improvements over the established technique of adaptive equalization to eliminate intersymbol interference and then making independent signal decisions in most cases did not turn out to be very significant, he turned his attention to using coding to improve performance. In this connection, it was clear to him that codes should be designed for maximum free Euclidean distance rather than Hamming distance, and that the redundancy necessary for coding would have to come from expanding the signal set to avoid bandwidth expansion.

To understand the potential improvements to be expected by this approach, he computed the channel capacity of channels with additive Gaussian noise for the case of discrete multilevel modulation at the channel input and unquantized signal observation at the channel output. The results of these calculations [2] allowed making two observations: firstly, that in principle coding gains of about 7-8 dB over conventional uncoded multilevel modulation should be achievable, and secondly, that most of the achievable coding gain could be obtained by expanding the signal sets used for uncoded modulation only by the factor of two. The author then concentrated his efforts on finding trellis-based signaling schemes that use signal sets of size 2^{m+1} for transmission of m bits per modulation interval. This direction turned out to be succesful and today's TCM schemes still follow this approach.

The next two sections illustrate with two examples how TCM schemes work. Whenever distances are discussed, Euclidean distances are meant.

Four-State Trellis Code for 8-PSK Modulation

The coded 8-PSK scheme described in this section was the first TCM scheme found by the author in 1975 with a significant coding gain over uncoded modulation. It was designed in a heuristic manner, like other simple TCM systems shortly thereafter. Figure 2 depicts signal sets and state-transition (trellis) diagrams for a) uncoded 4-PSK modulation and b) coded 8-PSK modulation with four trellis states. A trivial one-state trellis diagram is shown in Fig. 2a only to illustrate uncoded 4-PSK from the viewpoint of TCM. Every connected path through a trellis in Fig. 2 represents an allowed signal sequence. In

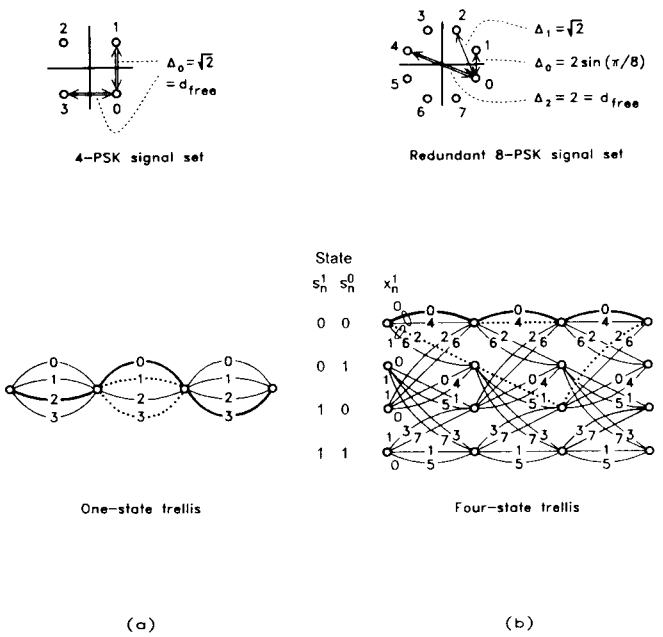

Fig. 2. (a) Uncoded four-phase modulation (4-PSK), (b) Four-state trellis-coded eight-phase modulation (8-PSK).

both systems, starting from any state, four transitions can occur, as required to encode two information bits per modulation interval (2 bit/sec/Hz). For the following discussion, the specific encoding of information bits into signals is not important.

The four "parallel" transitions in the one-state trellis diagram of Fig. 2a for uncoded 4-PSK do not restrict the sequences of 4-PSK signals that can be transmitted, that is, there is no sequence coding. Hence, the optimum decoder can make independent nearest-signal decisions for each noisy 4-PSK signal received. The smallest distance between the 4-PSK signals is $\sqrt{2}$, denoted as Δ_0. We call it the "free distance" of uncoded 4-PSK modulation to use common terminology with sequence-coded systems. Each 4-PSK signal has two nearest-neighbor signals at this distance.

In the four-state trellis of Fig. 2b for the coded 8-PSK scheme, the transitions occur in pairs of two parallel transitions. (A four-state code with four distinct transitions from each state to all successor states was also considered; however, the trellis as shown with parallel transitions permitted the achievement of a larger free distance.) Fig. 2b shows the numbering of the 8-PSK signals and relevant distances between these signals: $\Delta_0 = 2\sin(\pi/8)$, $\Delta_1 = \sqrt{2}$, and $\Delta_2 = 2$. The 8-PSK signals are assigned to the transitions in the four-state trellis in accordance with the following rules:

a) Parallel transitions are associated with signals with maximum distance Δ_2(8-PSK) = 2 between them, the signals in the subsets (0,4), (1,5), (2,6), or (3,7).
b) Four transitions originating from or merging in one state are labeled with signals with at least distance Δ_1(8-PSK) = $\sqrt{2}$ between them, that is, the signals in the subsets (0,4,2,6) or (1,5,3,7).
c) All 8-PSK signals are used in the trellis diagram with equal frequency.

Any two signal paths in the trellis of Fig. 2(b) that diverge in one state and remerge in another after more than one transition have at least squared distance $\Delta_1^2 + \Delta_0^2 + \Delta_1^2 = \Delta_2^2 + \Delta_0^2$ between them. For example, the paths with signals 0-0-0 and 2-1-2 have this distance. The distance between such paths is greater than the distance between the signals assigned to parallel transitions, Δ_2(8-PSK) = 2, which thus is found as the free distance in the four-state 8-PSK code: $d_{free} = 2$. Expressed in decibels, this amounts to an improvement of 3 dB over the minimum distance $\sqrt{2}$ between the signals of uncoded 4-PSK modulation. For any state transition along any coded 8-PSK sequence transmitted, there exists only one nearest-neighbor signal at free distance, which is the 180° rotated version of the transmitted signal. Hence, the code is invariant to a signal rotation by 180°, but to no other rotations (cf., Part II). Figure 3 illustrates one possible realization of an encoder-modulator for the four-state coded 8-PSK scheme.

Soft-decision decoding is accomplished in two steps: In the first step, called "subset decoding", within each subset of signals assigned to parallel transitions, the signal closest to the received channel output is determined. These signals are stored together with their squared distances from the channel output. In the second step, the Viterbi algorithm is used to find the signal path through the code trellis with the minimum sum of squared distances from the sequence of noisy channel outputs received. Only the signals already chosen by subset decoding are considered.

Tutorial descriptions of the Viterbi algorithm can be found in several textbooks, for example, [12]. The essential points are summarized here as follows: assume that the optimum signal paths from the infinite past to all trellis states at time n are known; the algorithm extends these paths iteratively from the states at time n to the states at time n + 1 by choosing one best path to each new state as a "survivor" and "forgetting" all other paths that cannot be extended as the best paths to the new states; looking backwards in time, the "surviving" paths tend to merge into the same "history path" at some time n − d; with a sufficient decoding delay D (so that the randomly changing value of d is highly likely to be smaller than D), the information associated with a transition on the common history path at time n − D can be selected for output.

Let the received signals be disturbed by uncorrelated Gaussian noise samples with variance σ^2 in each signal dimension. The probability that at any given time the decoder makes a wrong decision among the signals associated with parallel transitions, or starts to make a sequence of wrong decisions along some path diverging for more than one transition from the correct path, is called the error-event probability. At high signal-to-noise ratios, this probability is generally well approximated by

$$Pr(e) \simeq N_{free} \cdot Q[d_{free}/(2\sigma)],$$

where Q(.) represents the Gaussian error integral

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int^{\infty} exp(-y^2/2)dy,$$

and N_{free} denotes the (average) number of nearest-neighbor signal sequences with distance d_{free} that diverge at any state from a transmitted signal sequence, and remerge with it after one or more transitions. The above approximate formula expresses the fact that at high

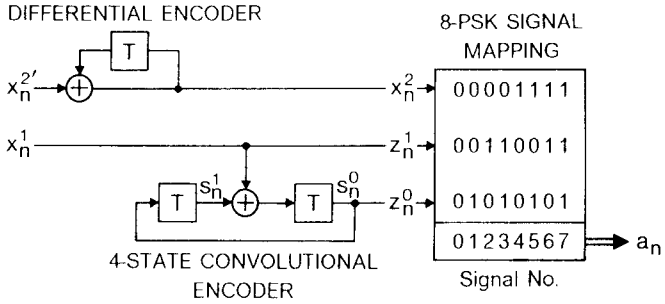

Fig. 3. Illustrates an encoder for the four-state 8-PSK code.

signal-to-noise ratios the probability of error events associated with a distance larger then d_{free} becomes negligible.

For uncoded 4-PSK, we have $d_{free} = \sqrt{2}$ and $N_{free} = 2$, and for four-state coded 8-PSK we found $d_{free} = 2$ and $N_{free} = 1$. Since in both systems free distance is found between parallel transitions, single signal-decision errors are the dominating error events. In the special case of these simple systems, the numbers of nearest neighbors do not depend on which particular signal sequence is transmitted.

Figure 4 shows the error-event probability of the two systems as a function of signal-to-noise ratio. For uncoded 4-PSK, the error-event probability is extremely well approximated by the last two equations above. For four-state coded 8-PSK, these equations provide a lower bound that is asymptotically achieved at high signal-to-noise ratios. Simulation results are included in Fig. 4 for the coded 8-PSK system to illustrate the effect of error events with distance larger than free distance, whose probability of occurrence is not negligible at low signal-to-noise ratios.

Figure 5 illustrates a noisy four-state coded 8-PSK signal as observed at complex baseband before sampling

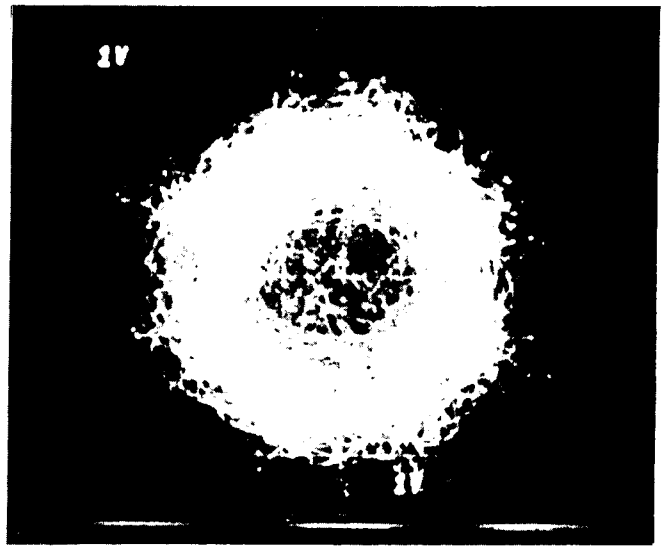

Fig. 5. Noisy four-state coded 8-PSK signal at complex baseband with a signal-to-noise ratio of $E_s/N_0 = 12.6$ dB.

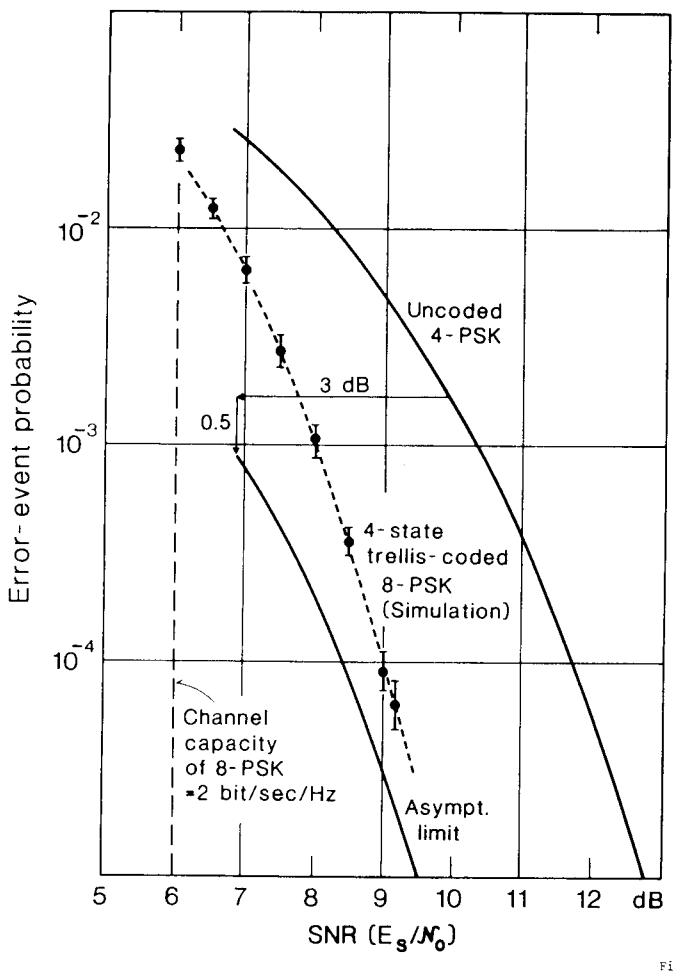

Fig. 4. Error-event probability versus signal-to-noise ratio for uncoded 4-PSK and four-state coded 8-PSK.

in the receiver of an experimental 64 kbit/s satellite modem [9]. At a signal-to-noise ratio of $E_s/N_0 = 12.6$ dB (E_s: signal energy, N_0: one-sided spectral noise density), the signal is decoded essentially error-free. At the same signal-to-noise ratio, the error rate with uncoded 4-PSK modulation would be around 10^{-5}.

In TCM schemes with more trellis states and other signal sets, d_{free} is not necessarily found between parallel transitions, and N_{free} will generally be an average number larger than one, as will be shown by the second example.

Eight-State Trellis Code for Amplitude/Phase Modulation

The eight-state trellis code discussed in this section was designed for two-dimensional signal sets whose signals are located on a quadratic grid, also known as a lattice of type "Z_2". The code can be used with all of the signal sets depicted in Fig. 1 for amplitude/phase modulation. To transmit m information bits per modulation interval, a signal set with 2^{m+1} signals is needed. Hence, for m = 3 the 16-QASK signal set is used, for m = 4 the 32-CROSS signal set, and soforth. For any m, a coding gain of approximately 4 dB is achieved over uncoded modulation.

Figure 6 illustrates a "set partitioning" of the 16-QASK and 32-CROSS signal sets into eight subsets. The partitioning of larger signal sets is done in the same way. The signal set chosen is denoted by A0, and its subsets by D0, D1, . . . D7. If the smallest distance among the signals in A0 is Δ_0, then among the signals in the union of the subsets D0,D4,D2,D6 or D1,D5,D3,D7 the minimum distance is $\sqrt{2}\ \Delta_0$, in the union of the subsets D0,D4; D2,D6; D1,D5; or D3,D7 it is $\sqrt{4}\ \Delta_0$, and within the individual subsets it is $\sqrt{8}\ \Delta_0$. (A conceptually similar partitioning of the 8-PSK signal set into smaller signal sets with increasing intra-set distances was implied in the example of coded 8-PSK. The fundamental importance

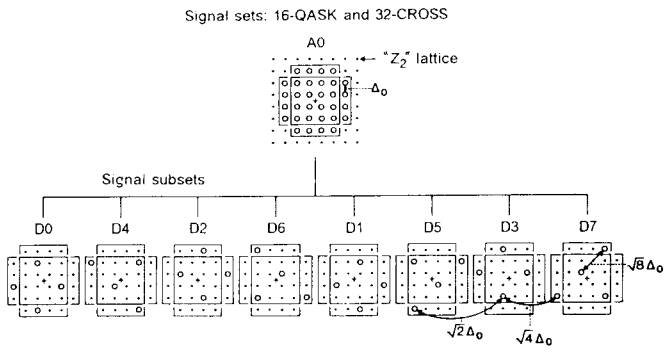

Fig. 6. Set partitioning of the 16-QASK and 32-CROSS signal sets.

of this partitioning for TCM codes will be explained in Part II.)

In the eight-state trellis depicted in Fig. 7, four transitions diverge from and merge into each state. To each transition, one of the subsets D0, . . . D7 is assigned. If A0 contains 2^{m+1} signals, each of its subsets will comprise 2^{m-2} signals. This means that the transitions shown in Fig. 7 in fact represent 2^{m-2} parallel transitions in the same sense as there were two parallel transitions in the coded 8-PSK scheme. Hence, 2^m signals can be sent from each state, as required to encode m bits per modulation interval.

The assignment of signal subsets to transitions satisfies the same three rules as discussed for coded 8-PSK, appropriately adapted to the present situation. The four transitions from or to the same state are always assigned either the subsets D0,D4,D2,D6 or D1,D5,D3,D7. This guarantees a squared signal distance of at least $2\Delta_0^2$ when sequences diverge and when they remerge. If paths remerge after two transitions, the squared signal distance is at least $4\Delta_0^2$ between the diverging transitions, and hence the total squared distance between such paths will be at least $6\Delta_0^2$. If paths remerge after three or more transitions, at least one intermediate transition contributes an additional squared signal distance Δ_0^2, so the squared distance between sequences is at least $\sqrt{5}\ \Delta_0$.

Hence, the free distance of this code is $\sqrt{5}\ \Delta_0$. This is smaller than the minimum signal distance within in the subsets D0, . . . D7, which is $\sqrt{8}\ \Delta_0$. For one particular code sequence D0-D0-D3-D6, Fig. 6 illustrates four error paths at distance $\sqrt{5}\ \Delta_0$ from that code sequence; all starting at the same state and remerging after three or four transitions. It can be shown that for any code

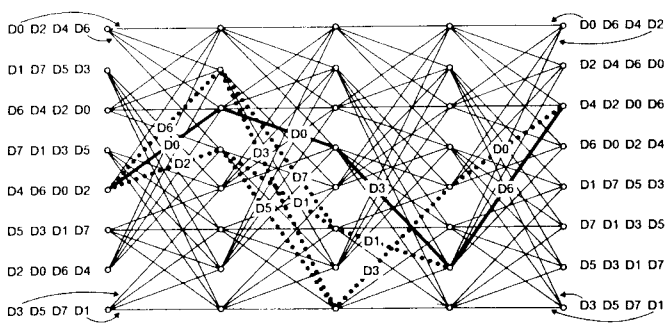

Fig. 7. Eight-state trellis code for amplitude/phase modulation with "Z_2"-type signal sets; $d_{free} = \sqrt{5}\ \Delta_0$.

sequence and from any state along this sequence, there are four such paths, two of length three and two of length four. The most likely error events will correspond to these error paths, and will result in bursts of decision errors of length three or four.

The coding gains asymptotically achieved at high signal-to-noise ratios are calculated in decibels by

$$G_{c,u} = 10\ log_{10}\ [(d^2_{free,c}/d^2_{free,u})/E_{s,c}/E_{s,u})],$$

where $d^2_{free,c}$ and $d^2_{free,u}$ are the squared free distances, and $E_{s,c}$ and $E_{s,u}$ denote the average signal energies of the coded and uncoded schemes, respectively. When the signal sets have the same minimum signal spacing Δ_0, $d^2_{free,c}/d^2_{free,u} = 5$, and $E_{s,c}/E_{s,u} \simeq 2$ for all relevant values of m. Hence, the coding gain is $10\ \log_{10}(5/2) \simeq 4$ dB.

The number of nearest neighbors depends on the sequence of signals transmitted, that is N_{free} represents an average number. This is easy to see for uncoded modulation, where signals in the center of a signal set have more nearest neighbors than the outer ones. For uncoded 16-QASK, N_{free} equals 3. For eight-state coded 16-QASK, N_{free} is around 3.75. In the limit of large "Z_2"-type signal sets, these values increase toward 4 and 16 for uncoded and eight-state coded systems, respectively.

Trellis Codes of Higher Complexity

Heuristic code design and checking of code properties by hand, as was done during the early phases of the development of TCM schemes, becomes infeasible for codes with many trellis states. Optimum codes must then be found by computer search, using knowledge of the general structure of TCM codes and an efficient method to determine free distance. The search technique should also include rules to reject codes with improper or equivalent distance properties without having to evaluate free distance.

In Part II, the principles of TCM code design are outlined, and tables of optimum TCM codes given for one-, two-, and higher-dimensional signal sets. TCM encoder/modulators are shown to exhibit the following general structure: (a) of the m bits to be transmitted per encoder/modulator operation, $\tilde{m} \leq m$ bits are expanded into $\tilde{m} + 1$ coded bits by a binary rate-$\tilde{m}/(\tilde{m}+1)$ convolutional encoder; (b) the $\tilde{m} + 1$ coded bits select one of $2^{\tilde{m}+1}$ subsets of a redundant 2^{m+1}-ary signal set; (c) the remaining $m-\tilde{m}$ bits determine one of $2^{m-\tilde{m}}$ signals within the selected subset.

New Ground Covered by Trellis-Coded Modulation

TCM schemes achieve significant coding gains at values of spectral efficiency for which efficient coded-modulation schemes were not previously known, that is, above and including 2 bit/sec/Hz. Figure 8 shows the free distances obtained by binary convolutional coding with 4-PSK modulation for spectral efficiencies smaller than 2 bit/sec/Hz, and by TCM schemes with two-dimensional signal sets for spectral efficiencies equal to or larger than 2 bit/sec/Hz. The free distances of uncoded modulation at the respective spectral effi-

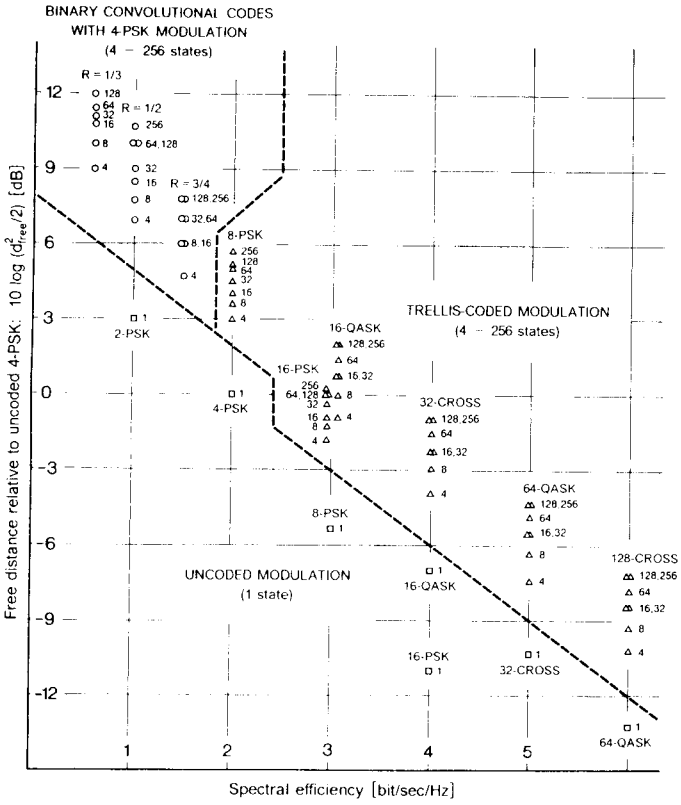

Fig. 8. Free distance of binary convolutional codes with 4-PSK modulation, and TCM with a variety of two-dimensional modulation schemes, for spectral efficiencies from 2/3 to 6 bit/sec/Hz.

ciencies are also depicted. The average signal energy of all signal sets is normalized to unity. Free distances are expressed in decibels relative to the value $d^2_{free} = 2$ of uncoded 4-PSK modulation. The binary convolutional codes of rates 1/3, 1/2, and 3/4 with optimum Hamming distances are taken from textbooks, such as, [12]. The TCM codes and their properties are found in the code tables presented in Part II (largely reproduced from [2]).

All coded systems achieve significant distance gains with as few as 4, 8, and 16 code states. Roughly speaking, it is possible to gain 3 dB with 4 states, 4 dB with 8 states, nearly 5 dB with 16 states, and up to 6 dB with 128 or more states. The gains obtained with two-state codes usually are very modest. With higher numbers of states, the incremental gains become smaller. Doubling the number of states does not always yield a code with larger free distance. Generally, limited distance growth and increasing numbers of nearest neighbors, and neighbors with next-larger distances, are the two mechanisms that prevent real coding gains from exceeding the ultimate limit set by channel capacity. This limit can be characterized by the signal-to-noise ratio at which the channel capacity of a modulation system with a 2^{m+1}-ary signal set equals m bit/sec/Hz [2] (see also Fig. 4).

Conclusion

Trellis-coded modulation was invented as a method to improve the noise immunity of digital transmission systems without bandwidth expansion or reduction of data rate. TCM extended the principles of convolutional coding to nonbinary modulation with signal sets of arbitrary size. It allows the achievement of coding gains of 3–6 dB at spectral efficiencies equal to or larger than 2 bit/sec/Hz. These are the values at which one wants to operate on many band-limited channels. Thus, a gap in the theory and practice of channel coding has been closed.

References

[1] 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.

[2] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Information Theory*, vol. IT-28, pp. 55–67, Jan. 1982.

[3] G. D. Forney, Jr., R. G. Gallager, G. R. Lang, F. M. Longstaff, and S. U. Qureshi, "Efficient modulation for band-limited channels," *IEEE Trans. Selected Areas in Comm.*, vol. SAC-2, pp. 632–647, Sept. 1984.

[4] L. F. Wei, "Rotationally invariant convolutional channel coding with expanded signal space—Part I: 180 degrees," *IEEE Trans. Selected Areas in Comm.*, vol. SAC-2, pp. 659–672, Sept. 1984.

[5] L. F. Wei, "Rotationally invariant convolutional channel coding with expanded signal space—Part II: nonlinear codes," *IEEE Trans. Selected Areas in Comm.*, vol. SAC-2, pp. 672–686, Sept. 1984.

[6] A. R. Calderbank and J. E. Mazo, "A new description of trellis codes," *IEEE Trans. Information Theory*, vol. IT-30, pp. 784–791, Nov. 1984.

[7] CCITT Study Group XVII, "Recommendation V.32 for a family of 2-wire, duplex modems operating on the general switched telephone network and on leased telephone-type circuits," Document AP VIII-43-E, May 1984.

[8] CCITT Study Group XVII, "Draft recommendation V.33 for 14400 bits per second modem standardized for use on point-to-point 4-wire leased telephone-type circuits," Circular No. 12, COM XVII/YS, Geneva, May 17, 1985.

[9] G. Ungerboeck, J. Hagenauer, and T. Abdel Nabi, "Coded 8-PSK experimental modem for the INTELSAT SCPC system," Proc. 7th Int. Conf. on Digital Satellite Communications (ICDS-7), pp. 299–304, Munich, May 12–16, 1986.

[10] R. J. F. Fang, "A coded 8-PSK system for 140-Mbit/s information rate transmission over 80-MHz nonlinear transponders," Proc. 7th Int. Conf. on Digital Satellite Communications (ICDS-7), pp. 305–313, Munich, May 12–16, 1986.

[11] T. Fujino, Y. Moritani, M. Miyake, K. Murakami, Y. Sakato, and H. Shiino, "A 120 Mbit/s 8PSK modem with soft-Viterbi decoding," Proc. 7th Int. Conf. on Digital Satellite Communications (ICDS-7), pp. 315–321, Munich, May 12–16, 1986.

[12] G. C. Clark and J. B. Cain, *Error-Correction Coding for Digital Communications*, Plenum Press, New York and London, 1981.

[13] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Information Theory*, vol. IT-13, pp. 260–269, April 1967.

[14] G. D. Forney, Jr., "The Viterbi algorithm," *Proc. of the IEEE*, vol. 61, pp. 268–278, March 1973.

[15] G. Ungerboeck, "Trellis-coded modulation with redundant signal sets, Part II: State of the art," *IEEE Communications Magazine*, vol. 25, no. 2, Feb. 1987.

Trellis-Coded Modulation with Redundant Signal Sets Part II: State of the Art

Gottfried Ungerboeck

This article is intended to bring the reader up to the state of the art in trellis-coded modulation. The general principles that have proven useful in code design are explained. The important effects of carrier-phase offset and phase invariance are discussed. Finally, recent work in trellis-coded modulation with multi-dimensional signal sets is described

February 1987—Vol. 25, No. 2
IEEE Communications Magazine

In this second part [1], a synopsis of the present state of the art in trellis-coded modulation (TCM) is given for the more interested reader. First, the general structure of TCM schemes and the principles of code construction are reviewed. Next, the effects of carrier-phase offset in carrier-modulated TCM systems are discussed. The topic is important, since TCM schemes turn out to be more sensitive to phase offset than uncoded modulation systems. Also, TCM schemes are generally not phase invariant to the same extent as their signal sets. Finally, recent advances in TCM schemes that use signal sets defined in more than two dimensions are described, and other work related to trellis-coded modulation is mentioned. The best codes currently known for one-, two-, four-, and eight-dimensional signal sets are given in an Appendix.

Design of Trellis-Coded Modulation Schemes

The trellis structure of the early hand-designed TCM schemes and the heuristic rules used to assign signals to trellis transitions suggested that TCM schemes should have an interpretation in terms of convolutional codes with a special signal mapping. This mapping should be based on grouping signals into subsets with large distance between the subset signals. Attempts to explain TCM schemes in this manner led to the general structure of TCM encoders/modulators depicted in Fig. 1. According to this figure, TCM signals are generated as follows: When m bits are to be transmitted per encoder/modulator operation, $\tilde{m} \leq m$ bits are expanded by a rate-$\tilde{m}/(\tilde{m}+1)$ binary convolutional encoder into $\tilde{m}+1$ coded bits. These bits are used to select one of $2^{\tilde{m}+1}$ subsets of a redundant $2^{\tilde{m}+1}$-ary signal set. The remaining $m-\tilde{m}$ uncoded bits determine which of the $2^{m-\tilde{m}}$ signals in this subset is to be transmitted.

Set Partitioning

The concept of set partitioning is of central significance for TCM schemes. Figure 2 shows this concept for a 32-CROSS signal set [1], a signal set of lattice type "Z_2". Generally, the notation "Z_k" is used to denote an infinite "lattice" of points in k-dimensional space with integer coordinates. Lattice-type signal sets are finite subsets of lattice points, which are centered around the origin and have a minimum spacing of Δ_0.

Set partitioning divides a signal set successively into smaller subsets with maximally increasing smallest intra-set distances Δ_i, $i = 0,1, \ldots$. Each partition is two-way. The partitioning is repeated $\tilde{m}+1$ times until $\Delta_{\tilde{m}+1}$ is equal to or greater than the desired free distance of the TCM scheme to be designed. The finally obtained subsets, labeled D0, D1, ... D7 in the case of Fig. 2, will henceforth be referred to as the "subsets." The labeling of branches in the partition tree by the $\tilde{m}+1$ coded bits $z_n^{\tilde{m}}, \ldots, z_n^0$, in the order as shown in Fig. 2, results in a label $z_n = [z_n^{\tilde{m}}, \ldots z_n^0]$ for each subset. The label reflects the position of the subset in the tree.

This labeling leads to an important property. If the labels of two subsets agree in the last q positions, but not in the bit z_n^q, then the signals of the two subsets are

0163-6804/87/0002-0012 $01.00

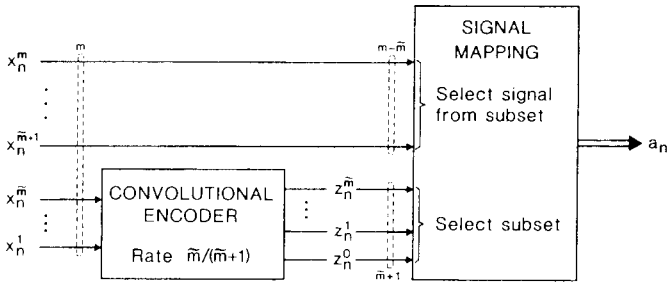

Fig. 1. General structure of encoder/modulator for trellis-coded modulation.

elements of the same subset at level q in the partition tree; thus they have at least distance Δ_q. This distance bound can be stated in a "set-partitioning lemma" and will be used in the next subsection.

The $m - \tilde{m}$ uncoded bits $x_n^m, \ldots, x_n^{\tilde{m}+1}$ are used to choose a signal from the selected subset. The specific labeling of subset signals by these bits is not particularly important at this point of the discussion. In the code trellis, the signals of the subsets become associated with $2^{m-\tilde{m}}$ parallel transitions.

The free Euclidean distance of a TCM code can now be expressed as

$$d_{free} = Min[\Delta_{\tilde{m}+1}, d_{free}(\tilde{m})],$$

where $\Delta_{\tilde{m}+1}$ is the minimum distance between parallel transitions and $d_{free}(\tilde{m})$ denotes the minimum distance between nonparallel paths in the TCM trellis diagram. In the special case of $\tilde{m} = m$, the subsets contain only one signal, and hence there are no parallel transitions.

Convolutional Codes for Trellis-Coded Modulation

At every time n, the rate-$\tilde{m}/(\tilde{m}+1)$ convolutional encoder depicted in Fig. 1 receives $\tilde{m}$ input bits, and generates $\tilde{m}+1$ coded bits which serve as the subset labels $\underline{z}_n = [z_n^{\tilde{m}}, \ldots z_n^0]$. The set of all possible sequences $\{\underline{z}_n\}$, which the encoder can generate, forms a convolutional code. A linear convolutional code of rate $\tilde{m}/(\tilde{m}+1)$ is most compactly defined by a parity-check equation which puts a constraint on the code bits in a sliding time window of length $\nu + 1$:

$$\sum_{i=0}^{\tilde{m}} (h_\nu^i z_{n-\nu}^i \oplus h_{\nu-1}^i z_{n-\nu+1}^i \oplus \ldots h_0^i z_n^i) = 0.$$

In this equation, $\oplus$ denotes modulo-2 addition. The quantity ν is called the constraint length. The quantities h_ℓ^i, $\nu \geq \ell \geq 0$; $0 \leq i \leq \tilde{m}$, are the binary parity-check coefficients of the code. Valid code sequences satisfy this equation at all times n. The equation defines only the code sequences, not the input/output relation of an encoder. A later subsection deals with minimal encoder realizations with ν binary storage elements, which is equivalent to saying that the code has 2^ν trellis states.

From the parity-check equation, one can observe that code sequences $\{\underline{z}_n\}$ can have arbitrary values for each $\tilde{m}$-tuple $[z_n^{\tilde{m}}, \ldots z_n^1]$ with an appropriate choice of the sequence $\{z_n^0\}$ so that the parity-check equation is satisfied. This property can be expressed in a "rate-$\tilde{m}/(\tilde{m}+1)$ code lemma."

Let now $\{\underline{z}_n\}$ and $\{\underline{z}'_n\} = \{\underline{z}_n \oplus \underline{e}_n\}$ be two code sequences, where $\{\underline{e}_n\}$ denotes the error sequence by which these sequences differ. Since the convolutional code is linear, $\{\underline{e}_n\}$ is also a code sequence. It follows from the "set-partitioning lemma" mentioned in the preceding subsection and the "rate-$\tilde{m}/(\tilde{m}+1)$ code lemma" that the squared free distance between non-parallel paths in the TCM trellis is bounded by [2]

$$d_{free}^2(\tilde{m}) \geq \underset{\{\underline{e}_n\} \neq \{0\}}{Min} \sum_n \Delta_{q(\underline{e}_n)}^2.$$

Here $q(\underline{e}_n)$ is the number of trailing zeros in $\underline{e}_n$, that is, the number of trailing positions in which two subset labels $\underline{z}_n$ and $\underline{z}'_n = \underline{z}_n \oplus \underline{e}_n$ agree. For example, $q(\underline{e}_n) = 2$, if $\underline{e}_n = [e_n^m, \ldots, e_n^3, 1, 0, 0]$. The "set-partitioning lemma" states that the distance between signals in the subsets selected by $\underline{z}_n$ and $\underline{z}'_n$ is lower-bounded by $\Delta_{q(\underline{e}_n)}$. One must take $\Delta_{q(\underline{0})} = 0$, not $\Delta_{\tilde{m}+1}$. Minimization has to be carried out over all non-zero code (error) sequences $\{\underline{e}_n\}$ that deviate at, say, time 0 from the all-zero sequence $\{\underline{0}\}$ and remerge with it at a later time. The "rate-$\tilde{m}/(\tilde{m}+1)$ code lemma" assures that for any given sequence $\{\underline{e}_n\}$ there exist two coded signal sequences whose signals have at any time n the smallest possible distance between the signals of subsets whose labels differ by $\underline{e}_n$. Usually, this smallest distance equals $\Delta_{q(\underline{e}_n)}$ for all $\underline{e}_n$. If this is the case, the above bound on $d_{free}(\tilde{m})$ becomes an equation. (Only when the signal subsets contain very few signals may the bound not be satisfied with equality. A similar always true equation can then be used to compute $d_{free}(\tilde{m})$ [2].)

This equation is of key importance in the search for optimum TCM codes. It states that free Euclidean distance can be determined in much the same way as free Hamming distance is found in linear binary codes, even though linearity does not hold for TCM signal sequences. It is only necessary to replace the Hamming weights of the e_n (number of 1's in $\underline{e}_n$) by the Euclidean weights $\Delta_{q(\underline{e}_n)}^2$. It is not necessary (as some authors seem to think) to compute distance between every pair of TCM signal sequences.

Search for Optimum TCM Codes

For the one- and two-dimensional signal sets depicted in Fig. 1 of Part I [1], the minimum intra-set distances are

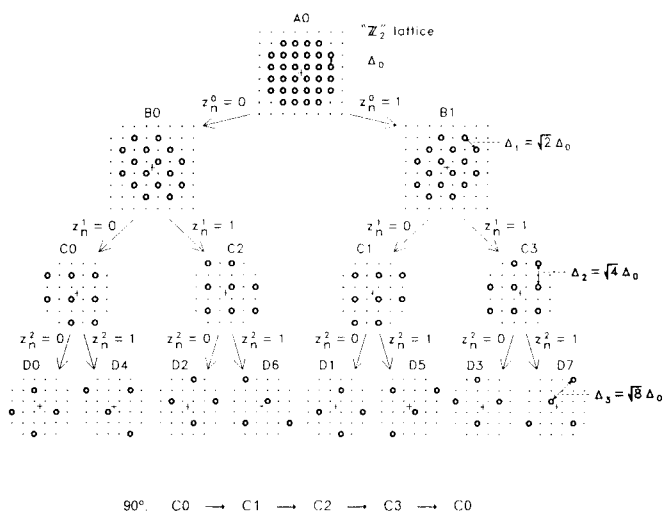

Fig. 2. Set partitioning of the 32-CROSS signal set (of lattice type "Z_2").

as follows. For 4-AM, 8-AM, ... (signal sets of type "Z_1"), $\Delta_{i+1} = 2\Delta_i$, $i = 0,1, \ldots$. For 16-QASK, 32-CROSS, ... (signal sets of lattice type "Z_2"), $\Delta_{i+1} = \sqrt{2}\,\Delta_i$, $i = 0,1, \ldots$. The non-lattice type signal sets 8-PSK and 16-PSK have special sequences of intra-set distances. The intra-set distances for higher-dimensional signal sets will be given when multi-dimensional TCM schemes are discussed later in this article.

For a given sequence of minimum intra-set distances $\Delta_0 \leq \Delta_1 \leq \ldots \Delta_{\tilde{m}}$, and a chosen value of ν, a convolutional code with the largest possible value of $d_{free}(\tilde{m})$ can be found by a code-search program described in [2]. The program performs the search for the $(\nu + 1) \cdot (\tilde{m} + 1)$ binary parity-check coefficients in a particular order and with a set of code-rejection rules such that explicit checks on the value of $d_{free}(\tilde{m})$ are very frequently avoided.

Tables of optimum codes for one-, two-, four-, and eight-dimensional TCM schemes are shown in the Appendix. Parity-check coefficients are specified in octal form, for example, $[h_\nu^0, \ldots, h_0^0] = [1,0,0,0,1,0,1]$ is written as $\underline{h}^0 = 105_8$. Equivalent codes in terms of free distance will be obtained if the parity-check coefficients of $\underline{h}^i$ are added modulo-2 to the coefficients of $\underline{h}^k$, for $i > k$ [2]. If $\Delta_i = \Delta_k$, $\underline{h}^i$ and $\underline{h}^k$ may also be interchanged. When in the code tables the free distance of a code is marked by an asterisk (*), $d_{free}(\tilde{m})$ exceeds $\Delta_{\tilde{m}+1}$, and hence the free distance occurs only between the subset signals assigned to parallel transitions. These schemes have the smallest numbers of nearest neighbors. For example, the 256-state code for "Z_1"-type signals has this property. For large values of m, this code attains a full 6 dB coding gain with only two nearest neighbors.

Two Encoder Realizations

The parity-check equation specifies only the convolutional code. Encoders for the same code can differ in the input/output relation which they realize. Figure 3 illustrates two encoders for the 8-state linear code specified in Tables II and III ($\nu = 3$) in the Appendix. One is called a systematic encoder with feedback, the other a feedback-free encoder. Both encoders are minimal, that is, they are realized with ν binary storage elements. The transformation of one minimal encoder into the other follows from the structural properties of convolutional codes described in [3]. With a systematic encoder, the input bits appear unchanged at the output. Therefore, a systematic encoder cannot generate a catastrophic code, i.e., a code with no distance increase between two trellis paths that remain distinct for an unbounded length. This is also true, although far from being obvious, for an equivalent minimal feedback-free encoder [3].

The forward and backward connections in the systematic encoder are specified by the parity-check coefficients of the code. All codes presented in the Appendix have $h_\nu^0 = h_0^0 = 1$. This guarantees the realizability of an encoder in the form shown in Fig. 3a. The reader familiar with recursive digital filters will see that the parity-check equation is used (almost directly) to compute the bit z_n^0 from the other uncoded bits. Furthermore, all codes have $h_\nu^i = h_0^i = 0$, for $i > 0$. This

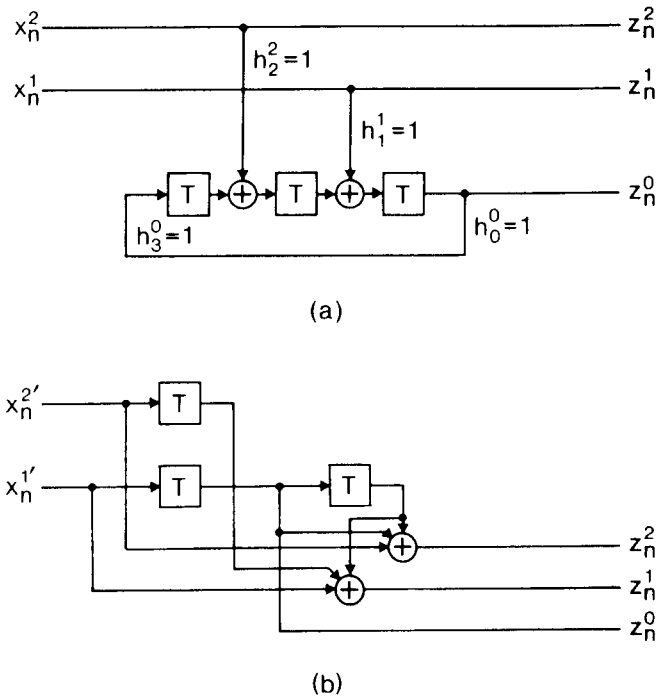

Fig. 3. Two encoders for a linear 8-state convolutional code with parity-check coefficients $\underline{h}^2 = [0,1,0,0]$, $\underline{h}^1 = [0,0,1,0]$, $\underline{h}^0 = [1,0,0,1]$ (cf. Tables II and III in the Appendix). (a) Minimal systematic encoder with feedback. (b) Minimal feedback-free encoder.

ensures that at time n the uncoded bits have no influence on the bit z_n^0, nor on the input to the first binary storage element in the encoder. Hence, whenever in the code trellis two paths diverge from or merge into a common state, the bit z_n^0 must be the same for these transitions, whereas the other bits differ in at least one bit. Signals associated with diverging and merging transitions therefore have at least distance Δ_1 between them, which reflects the second heuristic rule for good TCM codes mentioned in Part I [1].

TCM schemes for two-dimensional carrier modulation (with 8-PSK signal sets and "Z_2"-type signal sets) have up to the present time attracted the most attention. Practical realizations of these systems indicated that the effects of transmission impairments other than additive Gaussian noise on their performance need to be studied, in particular those of carrier offset.

Effects of Carrier-Phase Offset

This section addresses the problems that arise when a carrier-modulated two-dimensional TCM signal is demodulated with a phase offset $\Delta\phi$. The soft-decision decoder then operates on a sequence of complex-valued signals $\{r_n\} = \{a_n \cdot \exp(j\Delta\phi) + w_n\}$, where the a_n are transmitted TCM signals and the w_n denote additive Gaussian noise. The phase offset $\Delta\phi$ could be caused, for instance, by disturbances of the carrier phase of the received signal which the phase-tracking scheme of the receiver cannot track instantly.

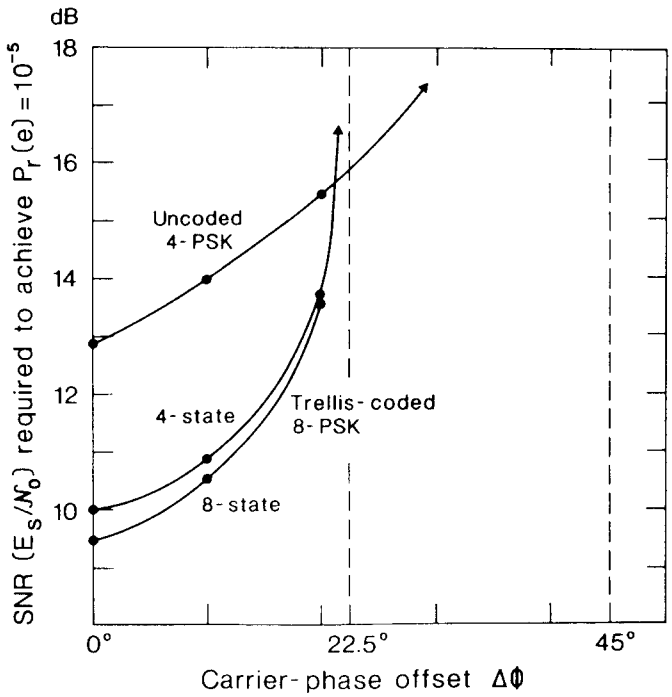

Fig. 4. Error performance of coded 8-PSK and uncoded 4-PSK in the presence of carrier-phase offset $\Delta\phi$.

Performance Degradation

The error performance of 4-state and 8-state coded 8-PSK systems in the presence of phase offset (based on unpublished work) is illustrated in Fig. 4. The figure shows the signal-to-noise ratio needed to sustain an error-event probability of 10^{-5} as a function of $\Delta\phi$. For the coded 8-PSK systems, the required signal-to-noise ratio increases with increasing values of $\Delta\phi$ until both systems fail at $\Delta\phi = 22.5°$, even in the absence of noise. In contrast, uncoded 4-PSK requires a higher signal-to-noise ratio at small phase offsets, but has an operating range up to $\Delta\phi = 45°$ in the absence of noise. These results are typical for TCM schemes.

The greater susceptibility of TCM schemes to phase offset can be explained as follows. In the trellis diagrams of TCM schemes, there exist long distinct paths with low growth of signal distance between them, that is, paths which have either the same signals or signals with smallest distance Δ_0 assigned to concurrent transitions. In the absence of phase offset, the non-zero squared distances Δ_0^2 and the squared larger distances of diverging or merging transitions add up to at least the squared free distance. However, if phase offset rotates the received signals such that received signals become located halfway between the signals of the original signal set, the difference in distance between received signals and the signals on distinct transitions that are Δ_0 apart may be reduced to zero. There may then be no difference in distance between a long segment of received signals and two distinct trellis paths, just as though the code were catastrophic. At this point, the decoder begins to fail.

Behavior of Carrier-Phase Tracking Loops

Nowadays, in most digital carrier-modulation systems, decision-directed loops are employed for carrier-phase tracking. In these loops, the phase offset is estimated from the received signal and the decoder decisions. The estimated phase offset controls the demodulating carrier phase. In a TCM receiver, if the phase offset exceeds a critical value, for example, 22.5° in the case of coded 8-PSK, the decoder decisions become essentially uncorrelated with the received signal and the mean value of the phase estimate drops to zero. Figure 5 illustrates, for 4-state coded 8-PSK [2], the mean estimate of $\Delta\phi$ ("S-curve") and its variance as a function of the actual value of the phase offset. A vanishing mean estimate, as occurs for $\Delta\phi$ between 22.5° and 157.5°, leaves the carrier-phase tracking loop in an undriven random-walk situation which can last for long periods. Eventually, the system resynchronizes when the randomly-fluctuating demodulating carrier phase approaches a value for which the received signal again resembles a valid TCM sequence. This behavior is in significant contrast to the short phase skips and rapid recovery observed in uncoded 4-PSK or 8-PSK systems. It suggests that in some cases TCM systems may require special methods to force rapid resynchronization.

Invariance of Two-Dimensional TCM Codes under Phase Rotation

TCM codes are not usually invariant to all phase rotations under which the signal set is phase invariant. Figure 5 indicates a phase symmetry of 4-state coded 8-PSK only at $\Delta\phi = 180°$, but not at other multiples of 45°. This symmetry can be verified by inspection of the code trellis presented in Fig. 2b of Part I [1]. Coded 8-PSK schemes which are invariant to phase shifts of all multiples of 45° have been found [4], but these schemes require more than four states to achieve a coding gain of 3 dB.

In general, it is desirable that TCM codes have as many phase symmetries as possible to ensure rapid carrier-phase resynchronization after temporary loss of synchronization. On the other hand, such phase invariances must be made transparent to the transmitted user

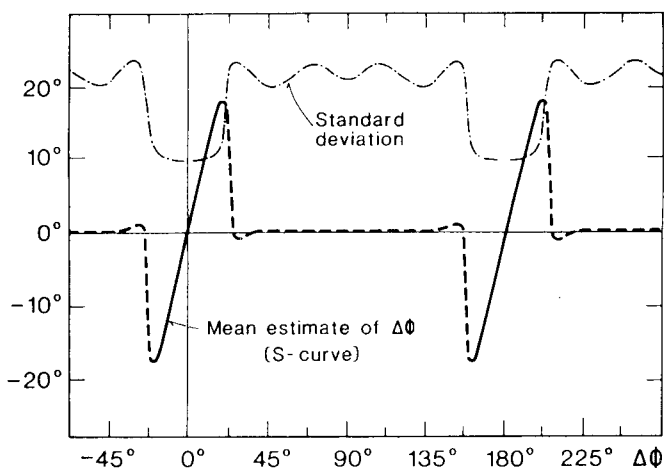

Fig. 5. Mean ("S-curve") and variance of the estimated phase offset $\Delta\hat{\phi}$ in a decision-directed carrier-phase tracking loop for 4-state coded 8-PSK versus the actual phase offset $\Delta\phi$, at a signal-to-noise ratio of 13 dB (tentative decisions used with zero delay).

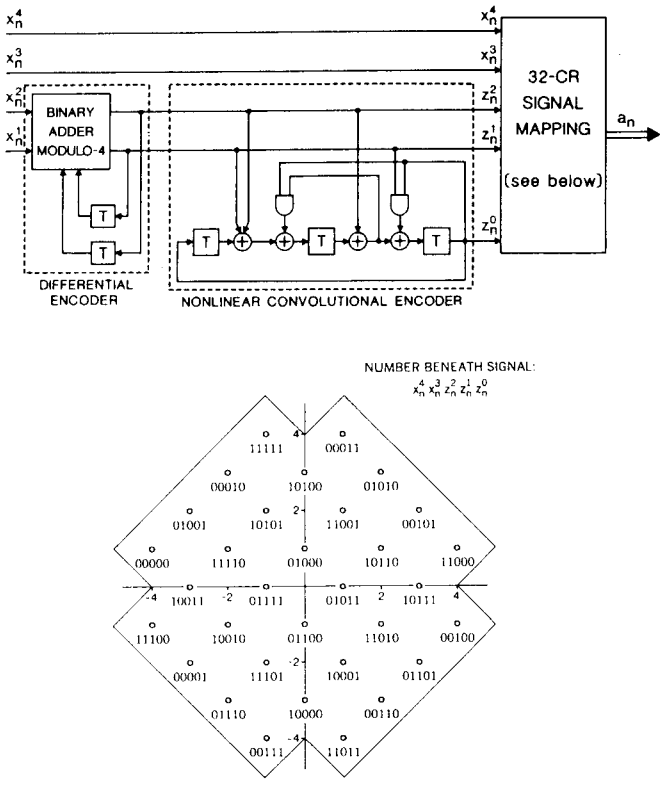

Fig. 6. Nonlinear 8-state encoder/modulator with 32-CROSS signal set and differential encoding, as in CCITT Recommendation V.32.

information by some form of differential encoding and decoding. If loss of phase synchronization is very unlikely, one may argue that TCM codes without phase invariances may have the advantage that the receiver can establish absolute phase from the received signal, so that no differential encoding/decoding is required.

The problems of phase invariance and differential encoding/decoding attracted considerable attention in work toward a TCM code for use in CCITT Recommendations for voice-band modems operating full-duplex at up to 9.6 kbit/s over two-wire telephone circuits, and at up to 14.4 kbit/s over four-wire circuits. There was considerable interest in a two-dimensional 8-state code that can achieve, with 90°-symmetric QASK and CROSS signal sets, a coding gain of about 4 dB over uncoded modulation. With the known linear code (cf. Table III in the Appendix, $\nu = 3$), it was only possible (by adding parity-check coefficients in a way which does not change free distance, as mentioned in the subsection on optimum-code search) to have either no phase symmetry or a symmetry at 180° [5], [4, Part I]. A breakthrough was finally accomplished by L.F. Wei, who introduced nonlinear elements into the convolutional encoder of the 8-state code. This made the code invariant to 90° rotations while maintaining its coding gain of 4 dB [6], [5, Part I]. Figure 6 shows the resulting encoder/modulator with its differential encoder, nonlinear convolutional encoder, and signal mapping for a 32-CROSS signal set (m = 4), as finally adopted in the CCITT V.32 Recommendation [7, Part I]. The labeling of subsets differs slightly from that indicated in Fig. 2, but the subsets are the same. The same code was also chosen for the CCITT V.33 Draft Recommendation [8, Part I], but with 64-QASK and 128-CROSS signal sets (m = 5,6). In the limit of large signal sets, the number of nearest neighbors in the 8-state linear and the CCITT nonlinear code is 16.

In a late contribution to the CCITT [7], illustrated in Fig. 7, an alternative 8-state nonlinear encoder with the differential-encoding function integrated into the encoder was proposed. The coding gain and the number of nearest neighbors are identical to those of the other 8-state schemes. The trellis diagram of the alternative nonlinear code was shown in Fig. 6 of Part I [1]. Differential decoding requires that the receiver compute $x_n^1 = z_n^0 \oplus z_{n+1}^0$. Subsets are labeled as indicated in Fig. 2. The selection of signals within the subsets by the uncoded bits x_n^1, x_n^3 is worth mentioning. If $x_n^1 = 0$, only signals of the inner 16-QASK set are transmitted (m = 3). With non-zero values of x_n^1, outer signals of the larger 32-CROSS set are also selected (m = 4). Extension of this concept to larger signal sets resulted in one general signal mapping for all data rates, e.g., for $3 \leq m \leq 7$ [7]. The mapping has the additional property that it can just as well be used for uncoded modulation with modulo-4 differential encoding of the bits z_n^1, z_n^0.

The nonlinear 8-state TCM codes appear to be special cases. Similar nonlinear phase-invariant codes with 16 and more states can be constructed. However, at least for 16 states, it does not seem possible to find a code with the same 4.8 dB coding gain as can be obtained with a linear code.

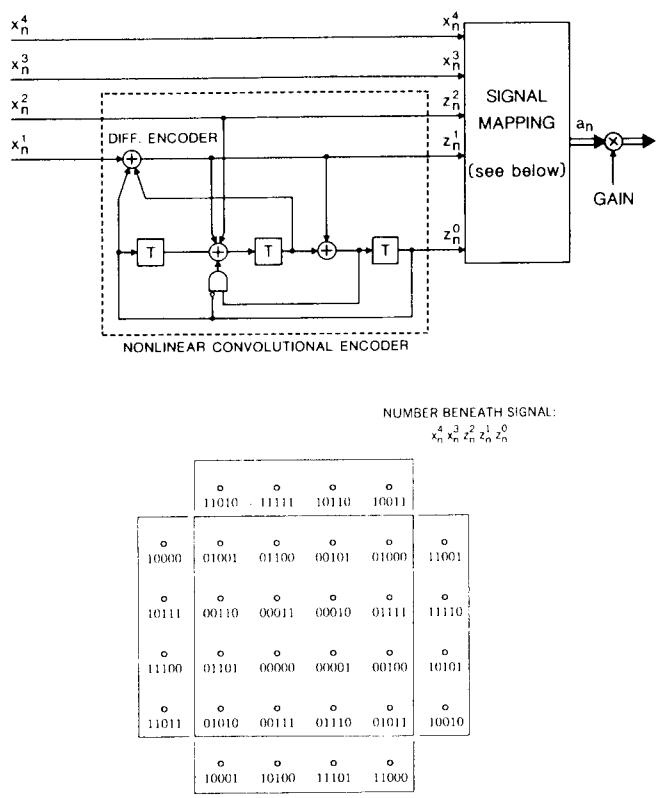

Fig. 7. Alternative nonlinear 8-state encoder modulator with integrated differential encoding and general signal mapping for 16-QASK, 32-CROSS, etc., signal sets.

Multi-Dimensional Trellis Codes

Recently, there have been a number of investigations into trellis coding with signal sets defined in more than two dimensions [3, Part I], [8–11]. In practical systems, multi-dimensional signals can be transmitted as sequences of constituent one- or two-dimensional (1-D or 2-D) signals. In this section, 2K-D TCM schemes are considered which transmit m bits per constituent 2-D signal, and hence mK bits per 2K-D signal. The principle of using a redundant signal set of twice the size needed for uncoded modulation is maintained. Thus, 2K-D TCM schemes use 2^{km+1}-ary sets of 2K-D signals. Compared to 2-D TCM schemes, this results in less signal redundancy in the constituent 2-D signal sets.

For 2-D TCM schemes with "Z_2"-type signal sets, the minimum signal spacing Δ_0 must be reduced by approximately the factor $\sqrt{2}$ (− 3 dB) to have the same average signal power as for uncoded modulation. This loss in signal spacing needs to be more than compensated for by coding to obtain an overall improvement in free distance. The lower signal redundancy of multi-dimensional TCM schemes with "Z_{2K}"-type signal sets results only in a reduction of the minimum signal spacing by the 2K-th root of 2 (−1.5 dB for K = 2; and −0.75 dB for K = 4), so coding has to contribute less than in the case of 2-D TCM to obtain the same gain in free distance. The larger signal spacing should also make multi-dimensional TCM systems less sensitive to phase offset. Finally, it has been found that multi-dimensional TCM schemes with 90° phase invariance can be obtained with linear codes.

Four-Dimensional Trellis-Coded Modulation

The 4-D TCM schemes (K = 2) described in this subsection employ compact sets of 2^{2m+1} signals chosen from a lattice of type "Z_4" with minimum signal spacing Δ_0. Figure 8 illustrates the set partitioning of a signal set A_1^0 of type "Z_4". The general idea is to derive the set partitioning of a higher-dimensional signal set from the set partitioning of constituent lower-dimensional signal sets. In the present case, A_1^0 and its subsets are characterized by two constituent "Z_2"-type signal sets A0 and their

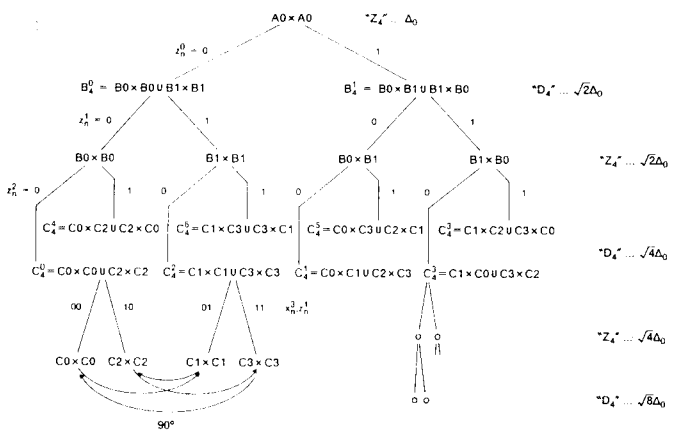

Fig. 8. Set partitioning of four-dimensional signal sets of lattice type "Z_4", also showing the effect of a 90° rotation.

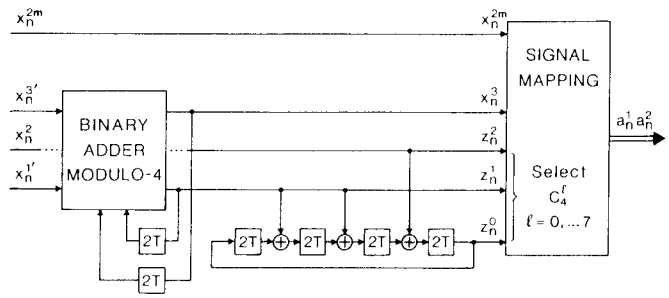

Fig. 9. Sixteen-state encoder/demodulator for four-dimensional "Z_4"-type trellis-coded modulation with differential encoding.

subsets, such as introduced in Fig. 2. This leads to a partition tree with signal sets of types "Z_4" → "D_4" → "Z_4" → "D_4" → "Z_4", etc., with minimum intra-set distances Δ_0, $\Delta_1 = \Delta_2 = \sqrt{2}\,\Delta_0$, $\Delta_3 = \Delta_4 = \sqrt{4}\,\Delta_0$, etc. The next paragraph describes the details of the partitioning process (and may be skipped by readers without specific interest in this process).

Set partitioning begins by writing $A_1^0 = A0 \times A0$ ($\times$ denotes set-product operation: the product set consists of all concatenations of elements of the first set with the elements of the second set). Substitution of A0 = B0 ∪ B1 (∪ denotes set union) yields A_1^0 = (B0∪B1)×(B0∪B1) = (B0×B0)∪(B0×B1)∪(B1×B0)∪(B1×B1). The first partition divides A_1^0 into the two subsets B_1^0 = (B0×B0)∪(B1×B1) and B_1^1 = (B0×B1)∪(B1×B0). These subsets are of type "D_4", where "D_4" denotes the densest lattice known in 4-D space [12]. The minimum intra-set distance in B_1^0 and B_1^1 is $\sqrt{2}\,\Delta_0$, which is the minimum distance between constituent 2-D signals in B0 or B1, and also between one 4-D signal in B0×B0 and another in B1×B1. On the next binary partition, e.g., when B_1^0 is partitioned into subsets B0×B0 and B1×B1, no distance increase is obtained. These subsets are of type "Z_4", like A_1^0, from which they differ only in their orientation, position with respect to the origin, and scaling. Hence, their partitioning is conceptually similar to that of A_1^0. The minimum intra-set distance increases to $\sqrt{4}\,\Delta_0$ when, e.g., B0×B0 is split into subsets C_4^0 = (C0×C0)∪(C2×C2) and C_4^1 = (C0×C2)∪(C2×C0), which are now again of type "D_4".

Optimum convolutional codes are found by using the obtained sequence of minimum intra-set distances in the code-search program mentioned earlier. The codes and their asymptotic coding gains over uncoded modulation with "Z_2"-type signals are given in Table IV in the Appendix. The gains are valid for large signal sets which fill the same volume in signal space as the signal sets used for uncoded modulation. Thus, the comparison is made for the same average signal power and the same peak power of 2-D signals.

It may be helpful to discuss the 16-state code of Table IV, which achieves an asymptotic coding gain of 4.52 dB, in more detail. The code uses the eight 4-D subsets C_4^0, ... C_4^7 shown in Fig. 8, and has 64 distinct transitions in its trellis diagram. The only nearest-neighbor signals are those associated with parallel transitions, and their number at any transition is 24 (the number of nearest neighbors in a "D_4" lattice). Figure 9 depicts one possible realization of an encoder/modulator with differential

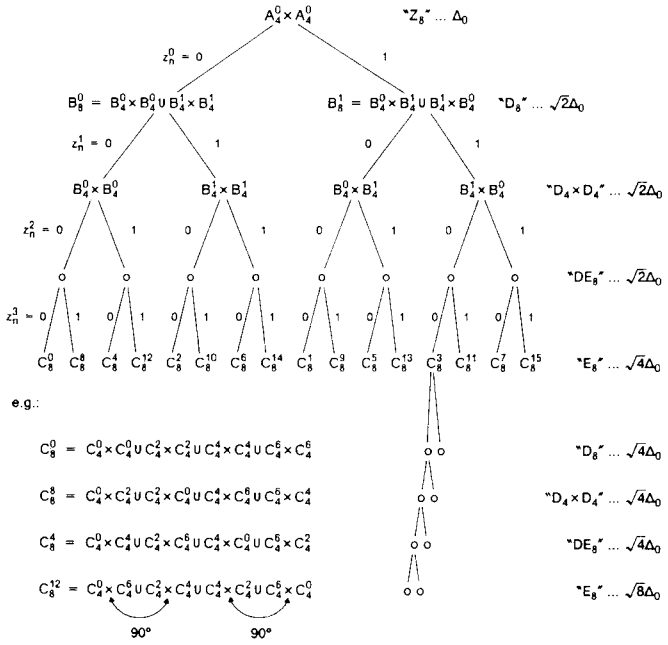

Fig. 10. Set partitioning of eight-dimensional signal sets of lattice type "Z_8", also showing the effect of a 90° rotation.

encoding. The code from Table IV was first made invariant to inversion of the bit z_n^1 by interchanging the parity-check coefficient vectors $\underline{h}^1$ and $\underline{h}^2$. Invariance to 90° rotations and the required differential encoding follow from the 90° symmetries indicated in Fig. 8, which in turn are based on the 90° symmetries in the constituent 2-D signal subsets. The subsets C_4^0, ... C_4^7, each composed of two subsets Ci×Ck, must be chosen individually for each value of m. The subset C0×C0 contains 2^{2m-3} signals, and may be constructed first. The other subsets Ci×Ck are then obtained by 90° rotations of the two constituent subsets C0 in C0×C0. For the specific case of m = 4.5, C0×C0 contains 8×8 signals, and hence the 8-ary subset C0 of Fig. 2 can be used. This construction of the 4-D subsets also suggests an efficient subset-decoding method that begins with signal decisions within the constituent 2-D subsets C0, ... C3. In general, the design of signal sets can be more complicated. References [3, Part I] and [11] discuss mapping techniques for cases where signal-set sizes are not powers of 2.

Eight-Dimensional Trellis-Coded Modulation

The technique of set partitioning of a higher-dimensional signal set based on the known partitioning of lower-dimensional sets is now applied to 8-D signal sets (K=4) of type "Z_8" = "$Z_4 \times Z_4$". Figure 10 illustrates the details. The sequence of minimum intra-set distances $\Delta_0, \Delta_1 = \Delta_2 = \Delta_3 = \sqrt{2}\Delta_0, \Delta_4 = \Delta_5 = \Delta_6 = \Delta_7 = \sqrt{4}\Delta_0$, etc., is obtained, corresponding to a chain of lattice types "Z_8" → "D_8" → "$D_4 \times D_4$" → "DE_8" → "E_8" → "D_8", etc., where "E_8" denotes the famous Gosset lattice, the densest lattice known in 8-D space [12]. (The nomenclature "DE_8" was introduced in [9]; [11] uses "D^4_8".)

Codes obtained by the code-search program are given in Table V in the Appendix. The codes use 2^{m+1} 8-D signals partitioned into 16 subsets C_8^0, ... C_8^{15} of type "E_8". In the limit of large signal sets, the codes achieve an asymptotic coding gain of 5.27 dB over uncoded "Z_2"-type modulation. If code complexity is increased to 64 states, the only nearest neighbors are those associated with parallel transitions, and their number is 240, which is the number of nearest neighbors in an "E_8" lattice. The "E_8"-type subsets can be further partitioned into two subsets with 90° symmetries as indicated in Fig. 10. This property can be verified by observing the 90° symmetries among the constituent 4-D signals as shown in Fig. 8. Hence, 8-D codes are inherently 90° phase invariant, because their subsets have this property. Differential encoding/decoding can be performed entirely within the subsets, decoupled from the convolutional encoding function.

Other 8-D TCM schemes are obtained by choosing the 2^{m+1} signals from another lattice type than "Z_8" in the chain of types encountered in Fig. 10, and performing the code search for the sequence of minimum intra-set distances that originates from this type. Codes with signals from "DE_8" or "E_8" are of some interest [9]-[11], although it does not seem that these codes exhibit significant advantages over the "Z_8"-type codes, if code complexities, asymptotic coding gains, and numbers of nearest neighbors are compared. This is also true for 4-D codes with "D_4" signals, as compared to codes with "Z_4" signals.

Discussion

The number of distinct transitions in the trellis diagrams of TCM codes is 2^{r+m}. This so-called "trellis complexity" represents a measure of code (decoding) complexity. A fair comparison of TCM schemes with different signal dimensionalities requires normalization of trellis complexities and numbers of nearest neighbors to the same number of signal dimensions. In the following, normalization to two dimensions is assumed. Hence, normalized trellis complexity specifies the number of distinct trellis transitions to be dealt with by the decoder per 2-D signal or two 1-D signals received. Similarly, a normalized number of nearest neighbors indicates the number of error events with free distance that could start (on average) during the same time interval.

In Fig. 11, asymptotic coding gains of TCM schemes with large 1-D (K = 0.5) to 8-D (K = 4) signal sets are plotted versus normalized trellis complexity, $2^{r+m}/K$. Normalized numbers of nearest neighbors, N_{free}/K, are given in parentheses. At a normalized trellis complexity of 8, the "Z_2"-type 4-state code is without competition. The "Z_4"-type 16-state code, whose encoder/modulator was illustrated in Fig. 9, shows a 0.5 dB advantage over a "Z_2"-type 8-state code, e.g., the nonlinear CCITT code, and also a slightly reduced number of nearest neighbors, at the same normalized complexity of 32. Next in the order of increasing complexities, the "Z_2"-type 16-state code may be of interest, but it cannot be made invariant to 90° rotations. At a normalized complexity of 128, i.e., four times the complexity of the CCITT code, the "Z_8"- and "E_8"-type 64-state codes are found as attractive 90° phase-invariant codes. Finally, at a 32 times higher complexity than the CCITT code, the "Z_4"-type 256-state code stands out for its asymptotic coding gain and low number of nearest neighbors.

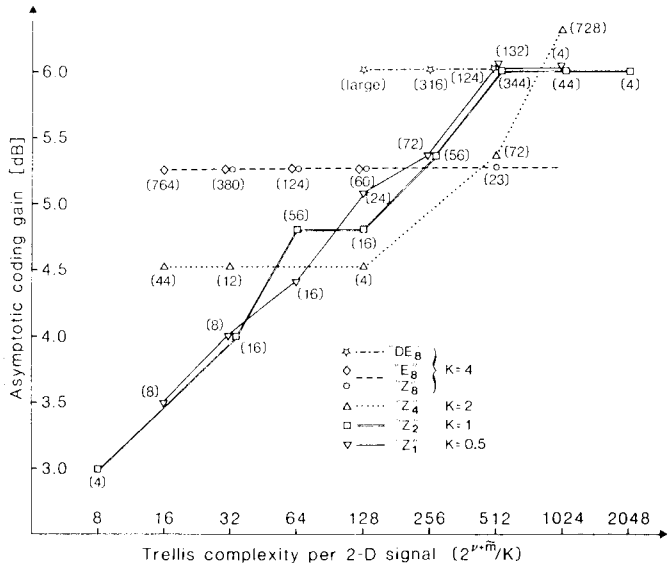

Fig. 11. Asymptotic coding gains versus trellis complexity per 2-D signal ($2^{\nu+\tilde{m}}/K$) for large 2K-D signal sets of type "Z_{2K}" for $K = 0.5, 1, 2, 4$; "E_8"; and "DE_8". Numbers of nearest neighbors per 2-D signal (N_{free}/K) are given in parentheses.

The asymptotic coding gain of the "DE_8" codes exceeds that of the "Z_8" and "E_8" codes by 0.75 dB, but the "DE_8" codes also have many more nearest neighbors. Hence, one may question their usefulness. Similarly, the "Z_1"-type 128-state code with the highest asymptotic coding gain of 6.28 dB shown in Fig. 11 may not be of practical interest, because of its large number of nearest neighbors.

Figure 11 gives important information about the ranking of TCM codes. However, the picture also remains somewhat incomplete. Real coding gains at given error probabilities, considering nearest and next-nearest neighbors and the boundary effect of finite signal sets, are not included. In first approximation, one may use the rule that for error rates around 10^{-5} the real coding gain is reduced by 0.2 dB for every increase in the number of nearest neighbors by the factor of 2. There is also very little published information about the carrier-phase sensitivity (a possible advantage of the multi-dimensional TCM schemes) of the TCM schemes under discussion. The complexity of subset decoding and decoder-memory requirements are further important aspects that need to be considered.

In general, one can make the following observations. At low complexity, higher-dimensional TCM schemes exhibit larger asymptotic coding gains than the lower-dimensional schemes, however, these coding gains are compromised by large numbers of nearest neighbors. In the mid-range, 4-D and 8-D TCM schemes achieve slightly larger real coding gains than the 1-D and 2-D schemes. Finally, at high trellis complexities lower-dimensional TCM schemes will eventually prevail in performance. This can be explained by the fact that these schemes have more signal redundancy available for coding than higher-dimensional TCM schemes. Overall, the differences in real coding gains are not very large, that is, they are smaller than 1 dB for the range of complexities considered.

Other Recent Work

Trellis codes have also been designed for 1-D and 2-D signal sets with nonequally-spaced ("asymmetric") signals [6, Part I], [13]. Some modest coding gains compared to schemes with equally-spaced signals are achieved when the codes have few states and small signal sets. These gains disappear for larger signal sets and higher code complexity. There are open questions about the number of nearest neighbors and sensitivity to carrier phase offset when signals are nonequally spaced.

While TCM schemes have been designed for linear modulation channels, similar developments took place in the field of continuous phase modulation (CPM) for channels requiring constant envelope signals. A summary on CPM schemes is given in [14].

Conclusion

It is probably fair to state that in recent years the theory of trellis-coded modulation has matured to the point where the achievement of further major gains seem less likely. However, there are still open questions concerning real coding gains, performance under channel impairments other than Gaussian noise, and actual implementation complexities.

The 8-state CCITT scheme was established only two years ago (1984). In the meanwhile, many manufacturers of voice-band modems and other transmission equipment have adopted the new combined coding and modulation technique. At least one manufacturer has already realized the sophisticated "Z_8"-type 64-state TCM scheme in a commercial product. In the struggle toward higher coding gains, application of more complexity is met with diminishing returns. For channels with Gaussian noise, the so-called "cut-off rate" R_0, which is smaller than channel capacity by the equivalent of about 3 dB, has been suggested as a more realistic limit [15]. TCM schemes have reached this barrier.

Acknowledgments

The author expresses his sincere appreciation for the many helpful comments and suggestions he obtained from colleagues and reviewers while writing this two-part article. Dr. G.D. Forney deserves special thanks for his many excellent technical comments and many detailed suggestions about the presentation of the material. By sending early versions of [9] and [11] to the author, Dr. L.F. Wei and Dr. D.G. Forney contributed significantly to the discussion of multi-dimensional TCM, as presented this part of the paper. Dr. V.M. Eyuboglu helped generously by verifying the correctness of the codes presented and providing still missing numbers of nearest neighbors.

Appendix: Code Tables

Tables I–III are largely reproduced from [2]. Tables IV and V have not been published previously; however, similar codes with up to 64 states were found by L.F. Wei [9]. In the tables, an asterisk (*) indicates that free distance occurs only among parallel transitions, i.e., $d_{free}(\tilde{m}) > \Delta_{\tilde{m}+1}$.

TABLE I
CODES FOR AMPLITUDE MODULATION WITH "Z_1" SIGNALS,
$\{\Delta_i, 0 \le i \le 2\} = \Delta_0, 2\Delta_0, 4\Delta_0$.

No. of states 2^ν	$\tilde{m}$	Parity check coefficients $\underline{h}^1$	$\underline{h}^0$	d^2_{free}/Δ_0^2	Asympt. coding gain [dB] $G_{4AM/2AM}$ (m = 1)	$G_{8AM/4AM}$ (m = 2)	$G_{C/u}$ (m→∞)	N_{free} (m→∞)
4	1	2	5	9.0	2.55	3.31	3.52	4
8	1	04	13	10.0	3.01	3.77	3.97	4
16	1	04	23	11.0	3.42	4.18	4.39	8
32	1	10	45	13.0	4.15	4.91	5.11	12
64	1	024	103	14.0	4.47	5.23	5.44	36
128	1	126	235	16.0	5.05	5.81	6.02	66
256	1	362	515	16.0*	—	5.81	6.02	2
256	1	362	515	17.0	5.30	—	—	

TABLE II
CODES FOR PHASE MODULATION
8-PSK: $\{\Delta_i, 0 \le i \le 2\} = 2\sin(\pi/8), \sqrt{2}, 2$;
16-PSK: $\{\Delta_i, 0 \le i \le 3\} = 2\sin(\pi/16), 2\sin(\pi/8), \sqrt{2}, 2$.

No. of states 2^ν	$\tilde{m}$	Parity-check coefficients $\underline{h}^2$	$\underline{h}^1$	$\underline{h}^0$	d^2_{free}/Δ_0^2	Asympt. coding gain [dB] $G_{8PSK/4PSK}$ (m=2)	$G_{16PSK/8PSK}$ (m=3)	N_{free} (m→∞)
4	1	—	2	5	4.000*	3.01	—	1
8	2	04	02	11	4.586	3.60	—	2
16	2	16	04	23	5.172	4.13	—	≈2.3
32	2	34	16	45	5.758	4.59	—	4
64	2	066	030	103	6.343	5.01	—	≈5.3
128	2	122	054	277	6.586	5.17	—	≈0.5
256	2	130	072	435	7.515	5.75	—	≈1.5
4	1	—	2	5	1.324	—	3.54	4
8	1	—	04	13	1.476	—	4.01	4
16	1	—	04	23	1.628	—	4.44	8
32	1	—	10	45	1.910	—	5.13	8
64	1	—	024	103	2,000*	—	5.33	2
128	1	—	024	203	2,000*	—	5.33	2
256	2	374	176	427	2.085	—	5.51	≈8.0

TABLE III
CODES FOR TWO-DIMENSIONAL MODULATION WITH "Z_2" SIGNALS,
$\{\Delta_i, 0 \le i \le 3\} = \Delta_0, \sqrt{2}\,\Delta_0, \sqrt{4}\,\Delta_0, \sqrt{8}\,\Delta_0$.

No. of states 2^ν	$\tilde{m}$	Parity-check coefficients $\underline{h}^2$	$\underline{h}^1$	$\underline{h}^0$	d^2_{free}/Δ_0^2	Asympt. coding gain [dB] $G_{16QA/8PSK}$ (m=3)	$G_{32CR/16QA}$ (m=4)	$G_{64QA/32CR}$ (m=5)	$G_{C/U}$ (m→∞)	N_{free} (m→∞)
4	1	—	2	5	4.0*	4.36	3.01	2.80	3.01	4
8	2	04	02	11	5.0	5.33	3.98	3.77	3.98	16
16	2	16	04	23	6.0	6.12	4.77	4.56	4.77	56
32	2	10	06	41	6.0	6.12	4.77	4.56	4.77	16
64	2	064	016	101	7.0	6.79	5.44	5.23	5.44	56
128	2	042	014	203	8.0	7.37	6.02	5.81	6.02	344
256	2	304	056	401	8.0	7.37	6.02	5.81	6.02	44
512	2	0510	0346	1001	8.0*	7.37	6.02	5.81	6.02	4

TABLE IV
CODES FOR FOUR-DIMENSIONAL MODULATION WITH "Z_4" SIGNALS,
$\{\Delta_i, 0 \le i \le 5\} = \Delta_0, \sqrt{2}\,\Delta_0, \sqrt{2}\,\Delta_0, \sqrt{4}\,\Delta_0, \sqrt{4}\,\Delta_0, \sqrt{8}\,\Delta_0$.

No. of states 2^ν	$\tilde{m}$	Parity-check coefficients $\underline{h}^4$	$\underline{h}^3$	$\underline{h}^2$	$\underline{h}^1$	$\underline{h}^0$	d^2_{free}/Δ_0^2	Asympt. coding gain [dB] (m→∞)	N_{free} (m→∞)
8	2	—	—	04	02	11	4.0	4.52	88
16	2	—	—	14	02	21	4.0*	4.52	24
32	3	—	30	14	02	41	4.0*	4.52	8
64	4	050	030	014	002	101	5.0	5.48	144
128	4	120	050	022	006	203	6.0	6.28	

TABLE V
CODES FOR EIGHT-DIMENSIONAL MODULATION WITH "Z_8" SIGNALS,
$\{\Delta_i, 0 \le i \le 5\} = \Delta_0, \sqrt{2}\,\Delta_0, \sqrt{2}\,\Delta_0, \sqrt{2}\,\Delta_0, \sqrt{4}\,\Delta_0, \sqrt{4}\,\Delta_0$.

No. of states 2^ν	$\tilde{m}$	Parity-check coefficients $\underline{h}^4$	$\underline{h}^3$	$\underline{h}^2$	$\underline{h}^1$	$\underline{h}^0$	d^2_{free}/Δ_0^2	Asympt. coding gain [dB] (m→∞)	N_{free} (m→∞)
16	3	—	10	04	02	21	4.0	5.27	
32	3	—	10	04	02	41	4.0	5.27	496
64	3	—	044	014	002	101	4.0*	5.27	240
128	4	120	044	014	002	201	4.0*	5.27	112

V.M. Eyuboglu and G.D. Forney [16] discovered typographical errors in the earlier published "Z_1"- and "Z_2"-type 256-state codes [2], which have now been corrected in Tables I and III.

Some of the 8-PSK codes of Table II were improved, compared to those published in [2], by using the exact expression for $d_{free}(\tilde{m})$ in the code search. The 16-PSK codes of Table II are new.

The exact numbers of nearest neighbors, N_{free}, given in the tables were taken from various sources, in particular [11] and [17]. The approximate values of N_{free}, given for some codes in Table II, are average values recently determined by the author.

References

[1] G. Ungerboeck, "Trellis-coded modulation with redundant signal sets—Part I: Introduction," *IEEE Communications Magazine*, vol. 25, no. 2, Feb. 1987.

[2] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Information Theory*, vol. IT-28, pp. 55-67, Jan. 1982.

[3] G. D. Forney, Jr., "Convolutional codes I: Algebraic structure," *IEEE Trans. Information Theory*, vol. IT-16, pp. 720-738, Nov. 1970.

[4] M. Oerder, "Rotationally invariant trellis codes for mPSK modulation," *1985 Internat. Commun. Conf. Record*, pp. 552-556, Chicago, June 23-26, 1985.

[5] IBM Europe, "Trellis-coded modulation schemes for use in data modems transmitting 3-7 bits per modulation interval," CCITT SG XVII Contribution COM XVII, No. D114, April 1983.

[6] AT&T Information Systems, "A trellis coded modulation scheme that includes differential encoding for 9600 bit/sec, full-duplex, two-wire modems," CCITT SG XVII Contribution COM XVII, No. D159, August 1983.

[7] IBM Europe, "Trellis-coded modulation schemes with 8-state systematic encoder and 90° symmetry for use in data modems transmitting 3-7 bits per modulation interval," CCITT SG XVII Contribution COM XVII, No. D180, October 1983.

[8] A. R. Calderbank and N. J. A. Sloane, "Four-dimensional modulation with an eight-state trellis code," *AT&T Tech. Jour.*, vol. 64, pp. 1005-1017, May-June 1985.

[9] L. F. Wei, "Trellis-coded modulation with multidimensional constellations," submitted to *IEEE Trans. Information Theory*, Aug. 1985.

[10] A. R. Calderbank and N. J. A. Sloane, "An eight-dimensional trellis code," *Proc. of the IEEE*, vol. 74, pp. 757-759, May 1986.

[11] G. D. Forney, Jr., *Coset Codes I: Geometry and Classification*, Aug. 25, 1986.

[12] N. J. A. Sloane, "The packing of spheres," *Scientific American*, vol. 250, pp. 116-125, Jan. 1984.

[13] M. K. Simon and D. Divsalar, "Combined trellis coding with asymmetric MPSK modulation," *JPL Publication* 85-24, May 1, 1985.

[14] C. E. Sundberg, "Continuous phase modulation," *IEEE Communications Magazine*, vol. 24, no. 4, pp. 25-38, April 1986.

[15] J. L. Massey, "Coding and modulation in digital communications," *Proc. 1974 Int. Zurich Seminar on Digital Communications*, Zurich, Switzerland, pp. E2(1)-(4), March 1974.

[16] V. M. Eyuboglu and G. D. Forney, Jr., private communications, Sept. 1984 and Sept. 1986.