Correctable error pattern	Syndrome
000000	000
000001	001
000010	010
000100	100
001000	101
010000	011
100000	110
100001	111

Suppose that the code word $\underline{c} = (110101)$ is transmitted and $\underline{r} = (010101)$ is received. We calculate the syndrome of $\underline{r}$ as:

received parity bits $= 101$
parity bits obtained by re-encoding received information $= \underline{011}$
$\therefore$ Syndrome $= 110$

From the table we note that (110) is the syndrome corresponding to the correctable error pattern $\underline{e} = (100000)$. Thus $\underline{r} + \underline{e} = (010101) + (100000) = (110101)$ is identified as the transmitted code word. Now suppose (011110) is transmitted and (101110) is received. The syndrome can be computed as before to obtain (101) which corresponds to (001000). The decoded word is identified as (101110) + (001000) = (100110). This is an incorrect decoding since the error pattern caused by the channel (110000) is not a correctable error pattern. This code corrects single errors in any position and one error pattern of double errors. Thus, as noted earlier, it is a single error correcting code.

Algebraic techniques—The most prominent among these is the iterative decoding algorithm for BCH codes due to Berlekamp. It is perhaps the deepest and most impressive theoretical result in coding theory (block or convolutional). A systems engineer who wishes to minimize the complexity of a BCH decoder is still well advised to use Berlekamp's procedure [3]. The algorithm was interpreted in terms of the design of a minimum-length shift register to produce a given sequence by Massey [2]. The key idea is to compute a so-called error-locator polynomial and solve for its roots. The complexity of this algorithm increases only as the square of the number of errors to be corrected. Thus, it is feasible to decode powerful codes. The use of Fourier-like transforms has recently been proposed as a vehicle for reducing decoder complexity [2].

The BCH decoder could be implemented at moderate data rates in a special purpose processor with an associated finite field arithmetic unit, and memory. Highly parallel realization has been used to achieve very high data rates (40 Mbps) [2].

The standard BCH decoding algorithm is a *bounded-distance* algorithm. That is, no error patterns of more than t errors can be corrected. This technique does not generalize easily to utilize soft decisions. At present, soft decisions can only be utilized via some other techniques in combination with the standard hard decision BCH decoding algorithm. Two such schemes are Forney's generalized minimum distance decoding and Chase's algorithm [2].

Permutation decoding is another example of algebraic decoding and the so-called error trapping is a special case of it [4]. This technique is based on the fact that if the weight of the syndrome for an (n,k) t-error correcting code is at most t, then the information bits are correct. If the weight of the syndrome is greater than t, then at least one information bit is incorrect.

Majority logic decoding—There are codes that, because of the special form of their parity check equations, are majority logic decodable. Majority-logic decoding is the simplest form of threshold decoding that is applicable to both block and convolutional codes. Recall that any syndrome bit is a linear combination of error bits. Thus, a syndrome bit represents a known sum of error bits. Further, any linear combination of syndrome bits is also a known sum of error bits. Hence, all 2^{n-k} such possible combinations of syndrome bits are of the known sum of error bits available at the receiver.

In the simplest case, decoding for these codes is performed on a bit-by-bit basis. For every received bit several parity check equations are checked giving rise to a particular value. The element 1 or 0 receiving the majority votes is taken to be the correct value for that bit. Many examples of this type of decoding procedure are given in [1].

Convolutional Codes and Their Decoding

Convolutional codes using either Viterbi or sequential decoding have the ability to utilize whatever soft-decision information might be available to the decoder. It is not surprising that they have been used widely even though their theory is not as mathematically profound as that of block codes. Most good convolutional codes have been found by computer search rather than by algebraic construction.

Convolutional codes can be studied from many diffferent approaches. For the purpose of illustrating their decoding methods, it is necessary to outline both the tree and the trellis approaches.

The convolutional encoder of Fig. 2 can be described by the code tree of Fig. 9. Each branch of the tree represents a single input bit—an input zero corresponds to the upper branch and an input one corresponds to the lower branch. Clearly any sequence of input bits traces out a particular path through the tree. Specifically, a 10110 sequence traces out a 11 01 00 10 10 output sequence.

In Fig. 9 we have labeled each node of the tree with a member from the following set of binary pairs: {00, 01, 10, 11} corresponding to the contents of the two left-most positions of the encoder register at that point in the tree. This number is called the state of the encoder.

We see that the tree contains redundant information which can be eliminated by merging, at any given level, all nodes corresponding to the same encoder state. The redrawing of the tree with merging paths has been called a trellis by Forney. Figure 10 represents a trellis for the convolutional encoder of Fig. 2. As before, an input 0 corresponds to the selection of the upper branch and an input 1 to the lower branch. Each

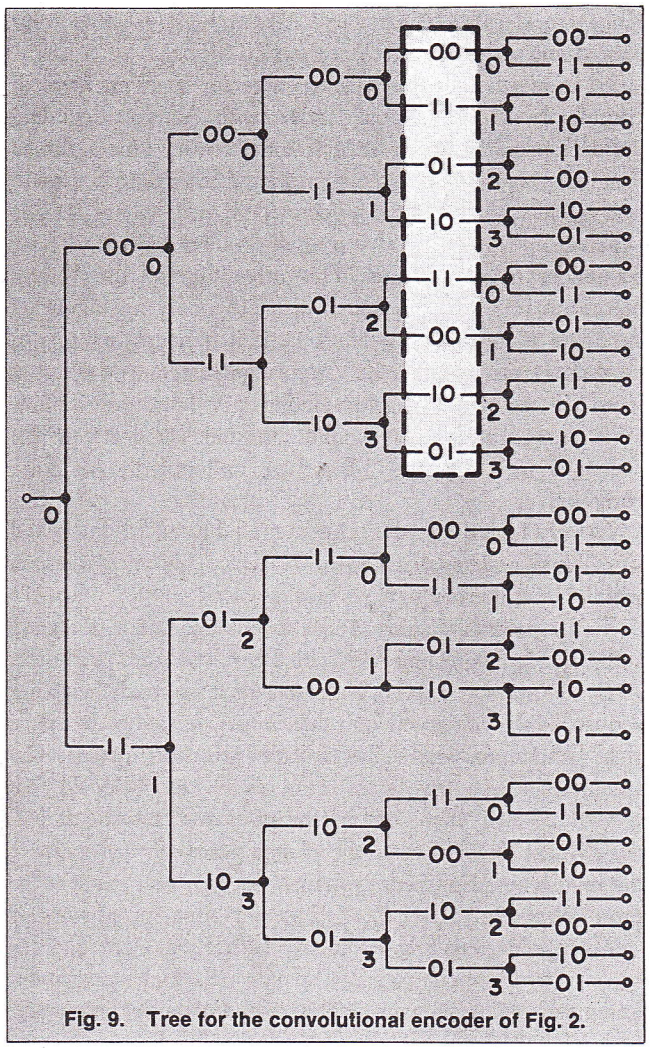

Fig. 9. Tree for the convolutional encoder of Fig. 2.

possible input sequence corresponds to a particular path in the trellis.

Unlike block codes, several distance measures have been proposed for convolutional codes, and each one is important and useful for particular decoding techniques.

The nth order column distance function $d_c(n)$ of a convolutional code is the minimum Hamming distance between all pairs of code words of length n branches which differ in their first branch of the code tree. The column distance function is a nondecreasing function of n, and assumes two particular values of special interest: d, the minimum distance of the code when $n = K$, the constraint length of the code; and d_f, the free distance of the code when $n \to \infty$.

The minimum distance of a convolutional code is the important parameter for determining the error probability of the code when used with threshold decoding. The free distance is useful in determining the code performance with Viterbi decoding and sequential decoding.

Decoding of Convolutional Codes

The problem of decoding a convolutional code can be thought of as attempting to find a path through the trellis or the tree by making use of some decoding rule.

Viterbi Decoding Algorithm

This algorithm leads to a maximum likelihood decoder for convolutional codes. In fact, it applies to any trellis code, not just convolutional codes. The significance of the trellis viewpoint is that the number of nodes in the trellis does not continue to grow as the number of input bits increases but remains at 2^{K-1}. The Viterbi algorithm computes a "metric" for every possible path through the trellis. It then discards a number of paths at every node that exactly balances the number of new paths that are created. Thus, it is possible to maintain a relatively small list of paths that are always guaranteed to contain the maximum-likelihood choice. The decoding algorithm can easily operate on soft-decisioned data. This is a major advantage of Viterbi decoding.

Viterbi decoding is presently the most important decoding technique for providing coding gain for a variety of channels. Unfortunately, it has been well over a decade since there have been any fundamentally new ideas in Viterbi decoding, and that technology appears to be near the asymptote of its learning curve [9].

We note that the complexity of the Viterbi algorithm is an exponential function of the code's constraint length K; unfortunately, the larger K is, the better the code is likely to be (that is, the larger are the coding gains that can be obtained). We are motivated to consider decoding algorithms that will work on convolutional codes with very large values of K, say $K >> 10$, the present limit of Viterbi decoders.

One last point worth mentioning is that Viterbi decoding does not perform very well in a bursty channel. In those channels, interleaving of data may thus have to be considered to obtain low correlation between noise samples. However, interleaving requires a significant increase in the encoding delay which may not be acceptable in certain applications.

Sequential Decoding

The complexity of sequential decoders is relatively independent of constraint length, so that codes with much larger constraint lengths can be used. A more rapid rate of change of error probability is achieved with increasing E_b/N_o. This technique is more suitable than Viterbi decoding when low bit error rates ($< 10^{-5}$) are required.

Sequential decoding was first introduced by Wozencraft but the most widely used algorithm to date is due to Fano. It is an efficient method for finding the most probable code word, given the received sequence, without searching the entire tree. The explored path is probably only local; that is, the procedure is suboptimum. The search is performed in a sequential manner, always operating on a single path, but the decoder can back up and change previous decisions. Each time the decoder moves forward, a "tentative" decision is made. If an incorrect decision is made, subsequent extensions of the path will be wrong. The decoder is eventually able to recognize this situation. When this happens, a substantial amount of computation is needed to recover the correct path. Backtracking and trying alternate paths continues until it finally decodes successfully.

A major problem with sequential decoding schemes is that

the number of computations required in advancing one node deeper into the code tree is so ill-behaved a random variable that even with very fast decoding circuitry and very large buffers, performance is limited by the probability of buffer overflow [9].

Threshold Decoding

Some convolutional codes are threshold decodable in that several parity checks are calculated for each message bit and if they exceed a threshold, a decision on the correctness of the bit is made. Moderate values of coding gain (1 to 3 dB) can be obtained with relatively inexpensive decoders and limited amount of redundancy.

Diffuse threshold decoding and the Gallager adaptive burst-finding scheme are two important variations of threshold decoding. These algorithms can also deal with burst errors [1].

Comparison of Block and Convolutional Codes

The theory of block codes is much older and richer than the theory of convolutional codes and the discussion on block codes is much longer than the discussion on convolutional codes. However, until recently this unbalance did not apply to practical applications. The discussion presented here is applicable to an additive Gaussian white noise channel. For spread spectrum systems with jamming and fading channels, the benefits of using codes (either block or convolutional) are even more spectacular.

We will take as our basis of comparison an uncoded BPSK system employing coherent detection. The same can be extended to QPSK, since the four-phase modulation may be considered as being the superposition of two BPSK systems each acting upon the orthogonal sine and cosine components of the carrier signal. We will assume that the information rate is fixed for all coded systems. The coded system will require more RF bandwidth. Comparisons among different techniques will be made at bit error rates = 10^{-5} and 10^{-8}. Table 1 has been adapted from [2]. Here the column labelled "data rate capability" is taken to be the following: low (less than 10 Kbps), moderate (10 Kbps to 1 Mbps), high (1 Mbps to 20 Mbps) and very high (greater than 20 Mbps).

At moderate and high data rates for a given level of complexity, convolutional codes with Viterbi decoding appears to be the most attractive technique. This assumes that there is no appreciable interference other than Gaussian noise; it also assumes that a decoded bit error rate of 10^{-5} is satisfactory and that the overall system transmits long sequences of bit streams. This advantage to the Viterbi algorithm follows because, in order to apply an algebraic decoding algorithm to a block code it is necessary to use hard-decisions, whereas the Viterbi algorithm can be adapted to accept soft-decisions with relative ease. However, if more efficient algorithms for decoding long block codes with soft decisions are developed, they will undoubtedly be quite competitive.

At very high data rates, concatenated Reed-Solomon and short block code systems can provide roughly the same gain with less complexity than Viterbi decoding.

For larger coding gains at high speeds, sequential decoding with hard decisions appears to be the most attractive choice. At moderate data rates a better case can be made for using sequential decoding with soft decisions.

In situations where the system protocols require the transmission of blocks of data (such as TDMA), all convolutional systems require flushouts and restarts, that is, the CODEC must be set to the all zero state before processing the next block. For such systems, block codes appear to be more attractive.

Threshold decoding is a very attractive technique for systems operating at very high speeds with small complexity. In digital satellite systems a potential use of this technique will be in digital telephony, where the user requires a smaller bit rate than that of the uncoded system but desires extremely high bandwidth efficiency.

For mobile terminals operating in the presence of large doppler offset and doppler rate, and multipath and fading, the use of Reed-Solomon codes with soft-decision decoding ap-

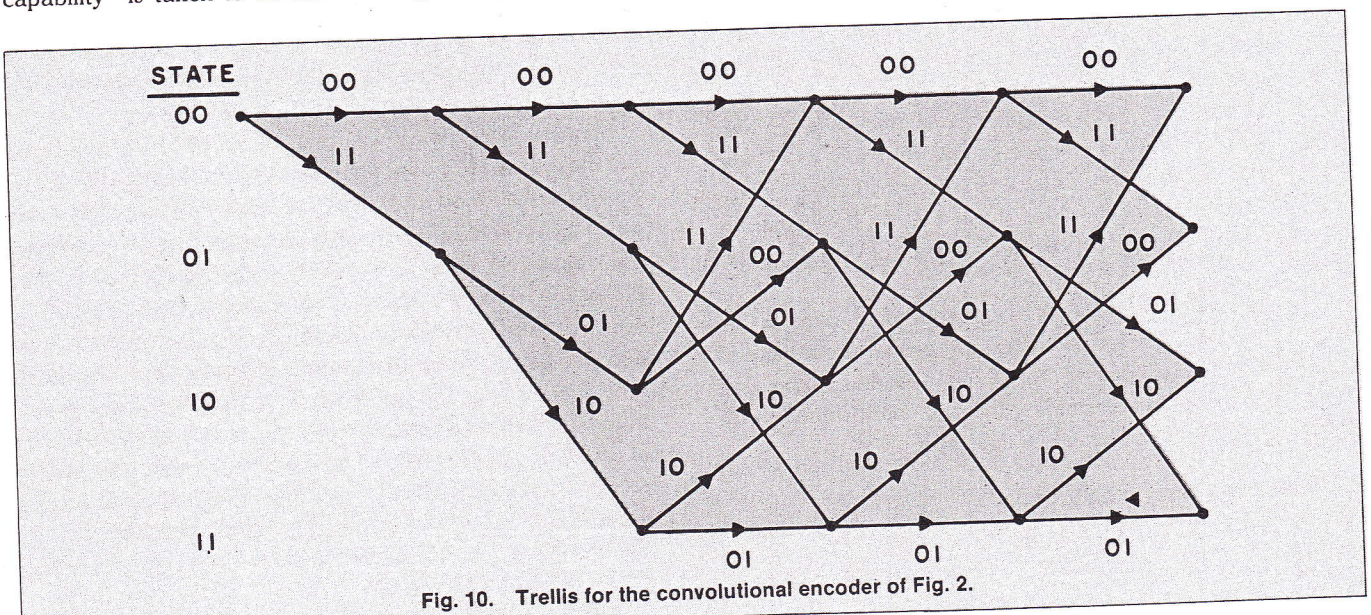

Fig. 10. Trellis for the convolutional encoder of Fig. 2.

TABLE 1
COMPARISON OF MAJOR CODING TECHNIQUES WITH BPSK OR QPSK MODULATION ON A GAUSSIAN CHANNEL

Coding technique	Coding gain(dB) at 10^{-5}	Coding gain(dB) at 10^{-8}	Data rate capability
Concatenated (RS and Viterbi)	6.5-7.5	8.5-9.5	Moderate
Sequential decoding (soft decisions)	6.0-7.0	8.0-9.0	Moderate
Block codes (soft decisions)	5.0-6.0	6.5-7.5	Moderate
Concatenated (RS and short block)	4.5-5.5	6.5-7.5	Very high
Viterbi decoding	4.0-5.5	5.0-6.5	High
Sequential decoding (hard decisions)	4.0-5.0	6.0-7.0	High
Block codes (hard decisions)	3.0-4.0	4.5-5.5	High
Block codes—threshold decoding	2.0-4.0	3.5-5.5	High
Convolutional codes—threshold decoding	1.5-3.0	2.5-4.0	Very high

pears extremely attractive to combat the bursty nature of the channel. Alternative coding techniques applicable are threshold decoders with interleaving or soft-decision Viterbi decoding with interleaving. However, interleaving requires a significant increase in the encoding delay. In packet networks where decoding, re-encoding and retransmission at several intermediate nodes is required, the delay associated with interleaved convolutional codes may not be acceptable. In such situations the solution might be to avoid interleaving by using one of the block codes suited for this purpose, which, rather than dispersing the bursts by interleaving, exploits it for improved error performance [9].

It should be stressed that a major part of these comparisons is influenced by today's digital integrated circuit technology. Advances in this technology could modify relative comparisons of complexity and achievable data rates.

Conclusion

The purpose of this paper was to introduce forward error correction schemes for digital communications. Various families of codes and their decoding methods were outlined. The performance of these codes over an additive Gaussian white noise channel was discussed.

An enormous amount of literature is now available on coding. For the interested reader, we have narrowed it down to a brief bibliography. More comprehensive lists of references are available in [1,2,4].

The theory of error correcting codes is an active area of research. Some critics claim that coding will be eliminated to conserve spectrum and space. This may not be the final answer since guard spaces, guard times, and minimum antenna separations are themselves users of spectrum, time, and space, and yet do not fully eliminate mutual interference and error [10]. Indeed, what Solomon Golomb wrote well over a decade ago is still very much true:

A message with content and clarity
Has gotten to be quite a rarity
To combat the terror
Of serious error
Use bits of appropriate parity.

Acknowledgment

The support of the author's research by the Natural Sciences and Engineering Research Council of Canada and by le Programme de Formation de Chercheurs et d'Action Concertée du Gouvernement du Quebec is gratefully acknowledged. The author would also like to thank Drs. Jean Conan, David Haccoun, and Gérald Séguin for helpful discussions.

Bibliography

[1] V. K. Bhargava, D. Haccoun, R. Matyas, and P. Nuspl, *Digital Communications by Satellite: Modulation, Multiple Access and Coding*, NY: Wiley, 1981.
[2] G. C. Clark, Jr. and J. B. Cain, *Error Correction Coding for Digital Communications*, NY: Plenum Press, 1981.
[3] R. J. McEliece, *The Theory of Information and Coding*, Reading, MA: Addison Wesley, 1977.
[4] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, Amsterdam, North-Holland and NY: Elsevier/North Holland, 1977.
[5] I. F. Blake and R. C. Mullin, *The Mathematical Theory of Coding*, NY: Academic Press, 1975.
[6] W. W. Peterson and E. J. Weldon, Jr., *Error Correcting Codes*, Second Edition, Cambridge, MA: MIT Press, 1972.
[7] S. Lin, *An Introduction to Error-Correcting Codes*, Englewood Cliffs, NJ: Prentice Hall, 1970.
[8] E. R. Berlekamp, *Algebraic Coding Theory*, NY: McGraw Hill, 1968.
[9] E. R. Berlekamp, "The technology of error-correcting codes," *Proc. IEEE*, vol. 68, pp. 564-593, May 1980.
[10] I. M. Jacobs, "Practical applications of coding," *IEEE Trans. Inf. Theory*, IT-20, pp. 305-310, May 1974.
[11] J. K. Wolf, "A survey of coding theory, 1967-1972," *IEEE Trans. Inf. Theory*, IT-19, pp. 381-389, July 1973.
[12] H. O. Burton and D. D. Sullivan, "Errors and error control," *Proc. IEEE*, vol. 60, pp. 1293-1301, Nov. 1972.
[13] R. T. Chien, "Block coding techniques for reliable data transmission," *IEEE Trans. on Commun. Technol.*, COM-19, pp. 743-751, Oct. 1971.
[14] A. J. Viterbi, "Convolutional codes and their performance in communication systems," *IEEE Trans. on Commun. Technol.*, COM-19, pp. 751-772, Oct. 1971.
[15] G. D. Forney, Jr., "Burst correcting codes for the classic bursty channel," *IEEE Trans. on Commun. Technol.*, COM-19, pp. 772-781, Oct. 1971.
[16] G. D. Forney, Jr., "Coding and its application in space communications," *IEEE Spectrum*, pp. 47-58, June 1970.
[17] J. F. Hayes, "The Viterbi algorithm applied to digital data transmission," *Communications Magazine*, vol. 13, no. 12, pp. 15-20, March 1975.

Vijay K. Bhargava was born in Beawar, India, on September 22, 1948. He received the B.Sc. (Math. and Eng.), M.Sc. (EE), and Ph.D. (EE) from Queen's University, Kingston, Ontario in 1970, 1972, and 1974 respectively.

He is currently an Associate Professor of Electrical Engineering at Concordia University in Montréal. His research interest is in the area of digital communications with special emphasis on error control coding techniques, cryptography and spread spectrum communications, and he has been a consultant to government agencies and industries in these areas. He is a coauthor of the book, *Digital Communications by Satellite*, published by John Wiley & Sons in December 1981. For the year 1982-1983 he is on sabbatical leave at Ecole Polytechnique de Montréal.

Vijay Bhargava is the junior past chairman of the IEEE Montréal Section and is a director of the IEEE Conferences Montréal, Inc. He is a co-vice chairman and the chairman (local arrangements) of the 1983 IEEE International Symposium on Information Theory. He was instrumental in the formation of the Information Theory Chapter in the Montréal Section and currently serves as its chairman. ■