DUAL-K CONVOLUTIONAL CODES FOR NONCOHERENTLY DEMODULATED CHANNELS

U. P. ODENWALDER
LINKABIT Corporation
San Diego, California

Summary. While the advantages of convolutional coding for coherently demodulated channels have become widely accepted, less work has been done on coding techniques for noncoherently demodulated channels used in channels experiencing fading or intentional interference. Here we describe a simple class of convolutional codes called dual-k [1] codes for use on 2^k-ary orthogonal signal modulated channels and show how they can be used with soft-decision Viterbi decoding on noncoherently demodulated channels.

The main result of this paper is a derivation of a closed form expression for the transfer function [2], T(D,N,L), for optimum (in the sense of best Hamming distance) dual-k convolutional codes. Examples of the technique of obtaining upper bounds on the bit error probability using this transfer function are also given for a noncoherently demodulated Rayleigh fading channel.

Introduction. Due to the greater implementation complexity of nonbinary convolutional coding systems, most systems which have been implemented have been binary. However, dual-k convolutional codes with 2^k-ary orthogonal signal modulation and soft-decision Viterbi decoding are simple enough for practical implementation in even the restricted space of a processing satellite.

Figure 1 shows a diagram of a rate 1/2 dual-3 encoder which has been hardware implemented. The input data is shifted in 3 bits at a time and then two octal outputs are generated. This code only has 8 binary states and can easily be decoded with a soft-decision Viterbi decoder. In fact, a more complex triple-3, i.e., a code with three octal registers and 64 binary states, would be feasible in many applications.

With Viterbi decoding the structure of the possible paths which can be compared with the all zero sequence path can be determined using the state diagram approach of Viterbi [2]. With this approach a state diagram is drawn with the all zero state split into an initial and a final state and the all zero state self loop omitted. The state transitions are labeled with terms of the form D^iN^jL where i represents the number of nonzero output characters produced by the transition, j represents the number of information character errors (0 or 1) produced by the transition, and L represents a path of length one branch.

The transfer function from the initial to the final all zero state represents all the paths which can be compared with the all zero path and the powers of D, N, and L represent the channel symbol errors, input character errors, and lengths, respectively, of the paths.

Here it will be shown that for rate R=1/v dual-k convolutional codes the code transfer function can be determined in a closed form. This closed form expression simplifies the performance analysis of these codes.

Transfer Function Derivation. A convenient way of describing dual-k convolutional codes is to treat the input and output symbols as 2^k-ary symbols over a finite field $GF(2^k)$ and to specify the code with $GF(2^k)$ tap multipliers. Figure 2 shows such a representation for a rate 1/v code. All the arithmetic in such a representation is over $GF(2^k)$.

With the finite field representation, the code words can be described by a generator matrix of the form

$$G=\begin{bmatrix} g_{01}\ g_{02}\cdots g_{0v} & g_{11}\ g_{12}\cdots g_{1v} & & \\ & g_{01}\ g_{02}\cdots g_{0v} & g_{11}\ g_{12}\cdots g_{1v} & \\ & & g_{01}\ g_{02}\cdots g_{0v} & g_{11} g_{12}\cdots g_{1v} \\ & & \ddots & \ddots \end{bmatrix} \tag{1}$$

The code words are linear combinations of the rows of this generator matrix. That is, if an L character information sequence is represented by the L dimensional row vector X, then the v(L+1) symbol output word represented by the row vector Y is given by

$$Y = XG \tag{2}$$

Before proceeding to the derivation of the transfer function of the code described by the generator matrix of (1), note from (2) that any common factor in the generator matrix may be factored out and the resulting set of code words has the same Hamming distance properties. The only difference between the set of code words is that the symbol of the words in one set are all multiplied by the common factor. In fact, since we are using the Hamming distance between paths where only the fact that the corresponding symbols are the same or not is of concern and since the only code parameters affecting the i-th symbol of a branch are the tap multipliers g_{0i} and g_{1i}, we may restrict our attention to the class of codes with a generator matrix of the form

$$G=\begin{bmatrix} 1\ \ 1\ \cdots 1 & g_{11}\ g_{12}\cdots g_{1v} & & \\ & 1\ \ 1\ \cdots 1 & g_{11}\ g_{12}\cdots g_{1v} & \\ & & 1\ \ 1\ \cdots 1 & g_{11}\ g_{12}\cdots g_{1v} \\ & & \ddots & \ddots \end{bmatrix} \quad (3)$$

when the g_{ij} are all nonzero. That is, when the g_{ij} are all nonzero, for any code specified by the generator matrix of (1) there exists an equivalent code (i.e., a code with the same distance properties) with a generator matrix of the form of (3).

From the generator matrix the minimum Hamming weight and thus the minimum distance between code words is seen to be 2v if none of the g_{1i} are zero. We will restrict our attention to this maximum free distance class of dual-k codes.

Another restriction we place on the class of dual-k codes is that the code not be catastrophic. That is, for an arbitrarily long information sequence, a finite number of channel errors should never produce an arbitrarily large number of decoded character errors. Massey and Sain [3] have given necessary and sufficient conditions for a convolutional code to be free of catastrophic error propagation. For the rate 1/v dual-k code this condition requires that the g_{1i} tap multipliers must not all be equal.

Summarizing, we will derive the transfer function for the class of noncatastrophic, maximum free distance, rate 1/vdual-k convolutional codes. This class of codes is equivalent to the class of codes specified by the generator matrix of (3) where none of the g_{1i} are zero and the g_{1i} are not all equal.

The transfer function of any noncatastrophic convolutional code is specified by the equations [2]

$$\begin{aligned} X &= AX + B \\ T(D, N, L) &= CX \end{aligned} \quad (4)$$

where X is a column vector whose components are equal to the transfer function from the all zero state to each of the nonzero states, A is a matrix whose components a_{ij} are the state or signal flow diagram branch transmission functions in going from the j-th to the i-th nonzero state, B is a column vector whose components are the branch transmission functions from the i-th nonzero state to the zero state, and C is a row vector whose components are the branch transmission functions from the zero state to the i-th nonzero state.

For the dual-k code considered here

$$B = D^v NL \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \tag{5}$$

and

$$C = D^V L \, [1 \;\; 1 \; \ldots \; 1] \tag{6}$$

The components of the A matrix depend on the particular code used. However, it has the property that for a particular code the sum of the components in any column of A is equal to a quantity, say α, which is independent of the column selected. Before proving this, note that this property enables us to write

$$CAX = \alpha CX \tag{7}$$

Thus, from (4) through (7)

$$T(D,\ N,\ L) = \frac{CB}{1-\alpha} = \frac{(2^k-1)D^{2v}L^2N}{1-\alpha} \tag{8}$$

To prove that the columns of A all sum to the same quantity consider the form of the branch output sequences. Referring to the generator matrix of (3), an input I and state S would produce a branch output

$$\text{Output Sequence} = \{I \oplus g_{11}S,\ I \oplus g_{12}S, \ldots,\ I \oplus g_{1v}S\} \tag{9}$$

Summing the columns of A is equivalent to summing the transmission functions corresponding to outputs of the form of (9) over all possible nonzero inputs. Note that I and S are nonzero and the g_{1i} have been restricted to be nonzero, Thus the only way for the i-th output symbol to be zero is for the input to equal $g_{1i}S$. So the a_{ij} transmission function is

$$a_{ij} = NLD^n \tag{10}$$

where n is the number of nonzero outputs on the branch transition.

To determine the sum of the column transmission functions first assume that all the g_{1i} are different. Then the sum is

$$\sum_{i=1}^{2^k-1} a_{ij}\Bigg|_{\substack{\text{all } g_{1i} \\ \text{different}}} = NL[vD^{v-1} + (2^k-1-v)D^v] \tag{11}$$

for $v \leq 2^k-1$. In general, if

β_1 of the g_{1i} are one nonzero value,
β_2 of the g_{1i} are another nonzero value,
.
.
.
β_n of the g_{1i} are another nonzero value

where

$\beta_1 + \beta_2 + ... + \beta_n = v \leq 2^k-1$

then

$$\alpha = \sum_{i=1}^{2^k-1} a_{ij} = NL\{D^{v-\beta_1} + D^{v-\beta_2} + ... + D^{v-\beta_n} + (2^k-1-n)D^v\} \tag{12}$$

Substituting (12) in (8) yields the transfer function.

Optimum Code Transfer Function. All of the codes considered here have the maximum free distance. So to select a code with the best distance properties we chose a code with the minimum number of words at the free distance plus 1. If more than one code has the same number of words at the free distance plus 1, we select a code in this class with the minimum number of words at the free distance plus 2. If more than one code remains, we continue examining the number of words at larger distances until a single code or a class of codes with the same distance properties remains. This code or class of codes is said to have the best distance properties.

By expanding the transfer function in a formal power series in D it is clear that the code with the best distance properties can be obtained by making all the g_{1i} different. The resulting transfer function is

$$T(D, N, L) = \frac{(2^k-1)D^{2v}L^2N}{1-NL[vD^{v-1} + (2^k-1-v)D^v]} \tag{13}$$

Sometimes it is of interest to know the transfer function of a R=1/v dual-k system in which every output symbol is repeated n times. The transfer function of such a system can be obtained by replacing D in (13) or,(8) and (12),by D^n.

Performance Bounds for the Independent-Symbol Rayleigh Fading Channel. To illustrate the application of this closed form transfer function, here we derive an upper bound on the bit error probability of a R=1/2 dual-3 coded 8-ary MFSK system on a Rayleigh fading channel.

Following Viterbi [2], let

$$\left.\frac{dT(D,N,L)}{dN}\right|_{\substack{N=1\\L=1}} = \frac{(2^k-1)D^{2v}}{[1-vD^{v-1}-(2^k-1-v)D^v]^2} = \sum_{i=2v}^{\infty} a_i D^i \tag{14}$$

The bit error probability can be upper bounded by

$$P_b < \frac{2^{k-1}}{2^k-1} \sum_{i=2v}^{\infty} a_i P_i \tag{15}$$

where P_i is the probability of an error in comparing two sequences which differ in i symbols. The $2^{k-1}/(2^k-1)$ factor converts the k-bit symbol error probability to bit error probability [4].

Many times since the first few terms are usually the only significant terms in the summation of (15), the bound can be tightened by using exact expressions or better bounds on the first few terms. We will illustrate this technique for the optimum R=1/2 dual-3 code.

Rewriting (14) for the R=1/2 dual-3 code gives

$$\left.\frac{dT(D,N,L)}{dN}\right|_{\substack{N=1\\L=1}} = 7\left\{D^4+4D^5+\frac{22D^6+4D^7-105D^8-100D^9}{(1-2D-5D^2)^2}\right\} \tag{16}$$

Then

$$P_b < 4P_4 + 16P_5 + \text{(tail terms)} \tag{17}$$

For an independent-symbol Rayleigh fading channel P_4 and P_5 can be bounded by [5, 6].

$$P_i \leq 2^{k-1} p^i \sum_{j=0}^{i-1} \binom{i-1+j}{j} (1-p)^j \tag{18}$$

where

$$p = \frac{1}{2 + \frac{k}{v} \frac{\bar{E}_b}{N_o}} \tag{19}$$

For the P_i contained in the tail terms of (17), we use the upper bound derived in the Appendix.

$$P_i < \frac{2^{k-1}}{\sqrt{5\pi}\;\; 2(1-2p)} [4p(1-p)]^i, \quad i \geq 6 \tag{20}$$

Then the tail term of (17) can be bounded by

$$\text{tail term} < (4)\frac{22D_0^6+4D_0^7-105D_0^8-100D_0^9}{\sqrt{5\pi}\;\; 2(1-2p)(1-2D_0-5D_0^2)^2} \tag{21}$$

where

$$D_0 = 4p(1-p) \tag{22}$$

and p is given by (19).

This bound is plotted in Figure 3. Figure 3 also gives upper bounds on the performance of 2 and 4 way diversity uncoded systems with octal signal modulation and noncoherent demodulation. For a diversity of L, i.e., L channel symbols per information symbol, the following uncoded bit error probability bounds were used

$$P_{uncoded}(L) < 2^{k-1} p_L^L \sum_{j=0}^{L-1} \binom{L-1+j}{j} (1-p_L)^j \tag{23}$$

where

$$p_L = \frac{1}{2 + \frac{k}{L} \frac{\bar{E}_b}{N_o}} \tag{24}$$

Comparing the curves in Figure 3 it can be seen that the coded system is much better than the uncoded L=2 system with the same number of channel symbols per information symbol and about 3 dB better than the uncoded L=4 system which has twice as many channel symbols per information symbol.

Conclusions. The closed form transfer function derived here simplifies the performance analysis of dual-k convolutional codes and the performance results obtained show that on a Rayleigh fading channel these codes can achieve a significant coding gain over uncoded systems with the same diversity.

Using techniques similar to those described here, performance bounds have also been obtained for dual-k convolutional coding systems in non-Gaussian environments. Using somewhat more sophisticated techniques the effects of receiver quantization have also been included in the bounds.

Appendix A. The probability of error for a binary orthogonal signal modulated noncoherently demodulated system with L-way diversity is [6]

$$\text{binary } P_b(L) = p^L \sum_{j=0}^{L-1} \binom{L-1+j}{j} (1-p)^j \qquad \text{(A.1)}$$

where

$$p = \frac{1}{2 + \frac{\bar{E}_s}{N_o}} \qquad \text{(A.2)}$$

Here we show that (A.1) can be upper bounded by

$$\text{binary } P_b(L) < \frac{[4p(1-p)]^L}{\sqrt{\pi(L-1)}\ 2(1-2p)}, \quad L > 1 \qquad \text{(A.3)}$$

For large L the bound of (A.3) is equal to (A.1).

To obtain the bound, first bound the binomial coefficient for L > 1

$$B_j = \binom{L-1+j}{j} \qquad \text{(A.4)}$$

Since

$$B_{j-1} = \frac{j}{L-1+j} B_j \leq \frac{1}{2} B_j, \quad j = 1, 2, \ldots, L-1 \qquad \text{(A.5)}$$

and [7]

$$B_{L-1} = \binom{2L-2}{L-1} < \frac{2^{2L-2}}{\sqrt{\pi(L-1)}} \tag{A.6}$$

We have

$$B_j < \frac{1}{2^{L-1-j}} \frac{2^{2L-2}}{\sqrt{\pi(L-1)}}, \qquad j = 0, 1, \ldots, L-1 \tag{A.7}$$

Substituting (A.7) into (A.1) and performing the sum yields

$$\text{binary } P_b(L) < \frac{(2p)^L}{2\sqrt{\pi(L-1)}} \sum_{j=0}^{L-1} [2(1-p)]^j$$

$$= \frac{(2p)^L}{2\sqrt{\pi(L-1)}} \frac{[2(1-p)]^L - 1}{2(1-p)-1}$$

$$< \frac{[4p(1-p)]^L}{2\sqrt{\pi(L-1)}\,(1-2p)} \tag{A.8}$$

This bound approaches the exact expression of (A.1) for large L.

The 2^k-ary symbol bound of (20) is obtained as in [5] by multiplying this bound by (2^k-1) and using the minimum L value (L=6 in this application) in the denominator term.

References

1. A. J. Viterbi, "Two Constructive Classes of Convolutional Codes for Multiple-Signal Channels," presented at the 1972 IEEE Intern. Symp. Inform. Theory, Pacific Grove, Calif., January 1972.

2. A. J. Viterbi, "Convolutional Codes and Their Performance in Communication Systems," IEEE Trans. Commun. Tech., Vol. COM-19, October 1971, pp. 751-772.

3. J. L. Massey and M. K. Sain, "Inverses of Linear Sequential Circuits," IEEE Trans. Computers, Vol. C-17, April 1968, pp. 330-337.

4. A. J. Viterbi, Principles of Coherent Communication. New York: McGraw-Hill, 1966.

5. A. J. Viterbi (ed), Advances in Communication Systems, Vol. 4. New York: Academic Press, 1975, pp. 279-308.

6. J. M. Wozencraft and I. M. Jacobs Principles of Communication Engineering. New York: Wiley, 1965.

7. R. M. Fano, The Transmission of Information. Cambridge, Massachusetts: M.I.T. Press, 1961, p. 216.

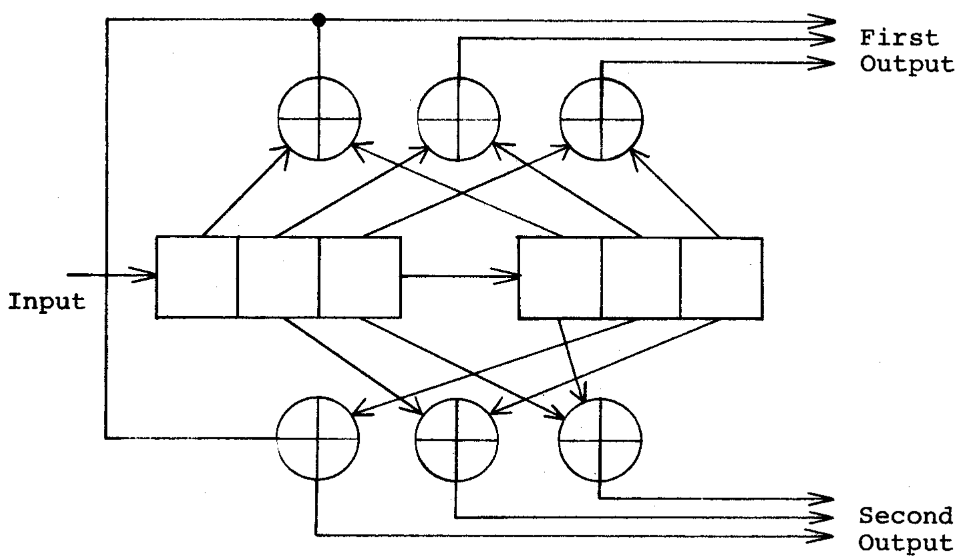

Figure 1. Rate 1/2 dual-3 convolutional encoder

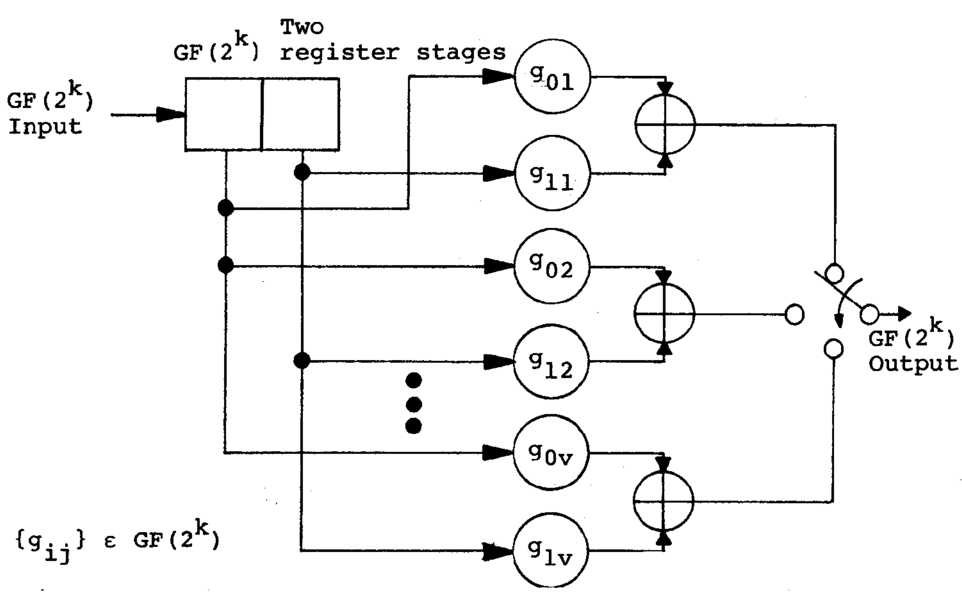

Figure 2. Finite field representation of a rate 1/v dual-k convolutional encoder

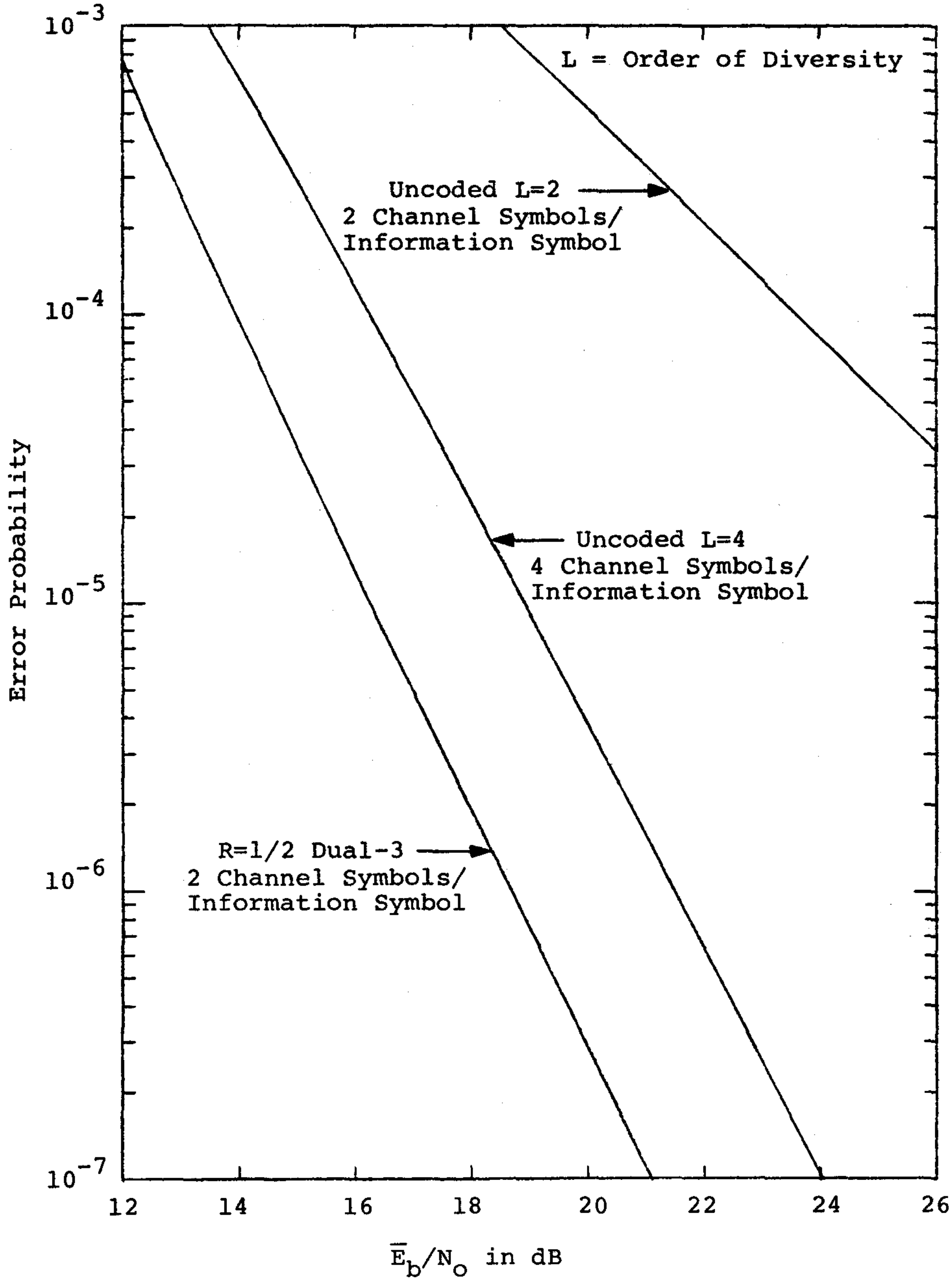

Figure 3. Performance of several noncoherently demodulated 8-ary MFSK systems on an independent Rayleigh fading channel