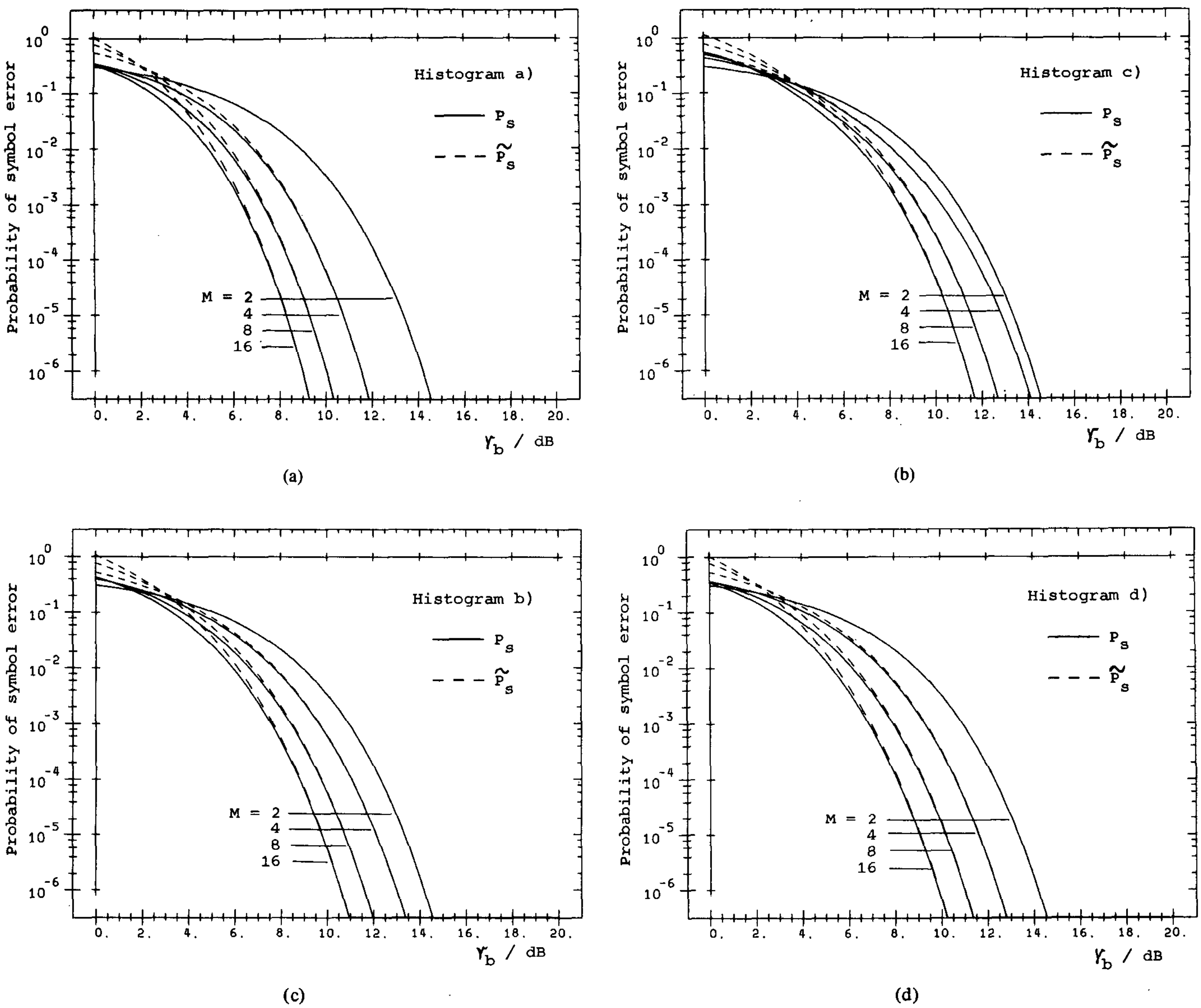


Fig. 3. Probability of error versus signal-to-noise ratio per information bit corresponding to the situation of interference as shown in Fig. 2(a)–(d).

Kaiserslautern, West Germany, for many valuable comments and suggestions. He also acknowledges helpful discussions with his colleague Dipl.-Ing. A. Baier at the University of Kaiserslautern.

REFERENCES

[1] S. Stein and J. J. Jones, *Modern Communication Principles.* New York: McGraw-Hill, 1967.

[2] P. W. Baier and M. Pandit, *Spread Spectrum Communication Systems* (Advances in Electronics and Electron Physics, vol. 53). New York: Academic, 1980, pp. 209–267.

[3] A. D. Whalen, *Detection of Signals in Noise.* New York: Academic, 1971.

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

[5] J. S. Bird, "Error performance of binary NCFSK in the presence of multiple tone interference and system noise," *IEEE Trans. Commun.*, vol. COM-33, pp. 203–209, Mar. 1985.

[6] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products.* New York: Academic, 1980.

Optimization of Trellis Codes with Multilevel Amplitude Modulation with Respect to an Error Probability Criterion

MICHAEL L. HONIG

Abstract—**An easily computed upper bound on the error probability of a data communications system with combined trellis coding and multilevel/phase modulation is derived, assuming an additive white Gaussian noise channel and maximum-likelihood decoding. This bound is used to search for codes obtained by set-partitioning that minimize the bound for a fixed number of trellis states. Only amplitude modulated signals**

Paper approved by the Editor for Modulation Theory and Nonlinear Channels of the IEEE Communications Society. Manuscript received April 25, 1985; revised January 17, 1986.

The author is with Bell Communications Research, Morristown, NJ 07960.

IEEE Log Number 8609576.

typically used in voiceband modem applications are considered. The signal levels that minimize the error probability bound subject to an average power constraint are presented for some specific codes.

I. Introduction

Binary convolutional coding combined with multilevel/phase modulation has been proposed as a means for reducing the signal-to-noise ratio (SNR) necessary to achieve a desired error rate in high speed data transmission. A rate m/n, $n > m$, multilevel trellis coder combines a rate m/n binary convolutional coder with a multilevel mapper, which maps the n output bits to one of 2^n points in a signal constellation, specifying in-phase and quadrature amplitudes. A union bound on the event error probability of a communications system with a multilevel trellis code, assuming an additive white Gaussian noise (AWGN) channel and Viterbi decoding, can be computed by using the error-state transition matrix [2], or the equivalent transfer function method [3], [4]. We show that if the mapping of coder output bits to signal constellation points is by set-partitioning [1], the computation required by this bound can be reduced substantially by eliminating the average over the transmitted source sequence. A search is performed to find codes that minimize the bound for a given SNR. Only amplitude modulated signals with one of eight levels (the 8-AM signal constellation) are considered, since this constellation has received much attention in voiceband data transmission applications [5], [6]. Nevertheless, consideration of other signal constellations presents no conceptual difficulties. The dependence of the union bound on the signal levels is also studied. Spacings between levels in 4-AM and 8-AM constellations that minimize the bound, assuming specific rate 1/2 and 2/3 codes, are found via a gradient search procedure.

II. Error Probability Bound for Multilevel Trellis Codes

A four-state, rate 2/3, 8-AM trellis code is shown in Fig. 1(a). The representation as a binary convolutional code and multilevel mapper is shown in Fig. 1(b). Without coding the two input bits would select one of the levels in the 4-AM constellation shown in Fig. 2(a). With coding the three output bits of the convolutional coder are mapped to a point in the 8-AM constellation, shown in Fig. 2(b). The symbols in Fig. 2 are assumed to be scaled to a peak transmitted power of unity. The uncoded bit in Fig. 1(b) corresponds to the parallel transitions shown in Fig. 1(a). Throughout the rest of this section, a rate $m/(m+1)$ trellis code is assumed in which the $m+1$ coded bits are mapped to a point in some one-dimensional signal constellation.

Assuming an AWGN channel, the output of the matched filter at the receiver at time T is

$$r_T = s_T + n_T \tag{1}$$

where s_T is the selected signal level at time T [i.e., a_j from Fig. 2(b)] and n_T is a Gaussian random variable.[1] (Consideration of channel impairments such as intersymbol interference and phase jitter is given in a related paper [7].) A maximum likelihood estimate of the transmitted symbols $s_1, s_2, \cdots, s_T$ given the received symbols $r_1, r_2, \cdots, r_T$ chooses the estimated sequence $\hat{s}_1, \hat{s}_2, \cdots, \hat{s}_T$ to minimize the sum

$$S = \sum_{i=1}^{T} (r_i - \hat{s}_i)^2 \tag{2}$$

over all possible estimates. This is usually accomplished via the Viterbi algorithm [3]. The Viterbi algorithm will choose

[1] Although one-dimensional signaling is assumed, this discussion is easily generalized to two dimensions by assuming that the variables in (1) are complex valued.

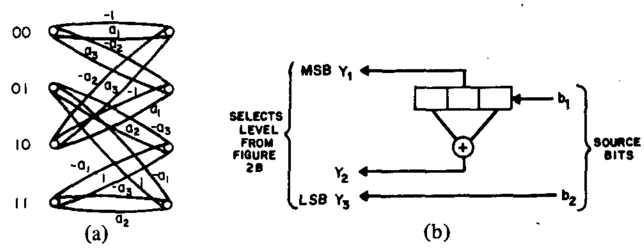


Fig. 1. Four-state, 8-AM trellis code. (a) Trellis diagram. (b) Convolutional code.

the incorrect path $P' = \{s_1', s_2', \cdots, s_T'\}$ rather than the transmitted path $P = \{s_1, s_2, \cdots, s_T\}$ if the metric of P' is smaller than the metric of P. Let $\Lambda(P)$ denote the metric of path $P = \{s_1, s_2, \cdots, s_T\}$, i.e.,

$$\Lambda(P) = \sum_{i=1}^{T} (r_i - s_i)^2. \tag{3}$$

Then the path metric difference is

$$\Lambda(P') - \Lambda(P) = D^2 + \tilde{n} \tag{4}$$

where $D^2 = \Sigma_{i=1}^{T} (s_i - s_i')^2$ and $\tilde{n} = 2\Sigma_{i=1}^{T} (s_i - s_i')n_i$. $\tilde{n}$ is a zero mean Gaussian random variable with variance $4\sigma^2 D^2$. The probability that the decoder will select path P' rather than path P is therefore

$$\text{Prob}\ (\Lambda(P) > \Lambda(P')) = \frac{1}{\sqrt{2\pi}} \int_{D/2\sigma}^{\infty} e^{-t^2/2}\, dt \le \frac{1}{2} e^{-D^2/8\sigma^2}. \tag{5}$$

The probability of at least one error event, given that path P_T of length T is transmitted starting at some arbitrary time in the trellis, is

$$\text{Prob}\left(\bigcup_{i=1}^{M_T} [\Lambda(P_{T,i}') < \Lambda(P_T)]\right) \le \sum_{i=1}^{M_T} \text{Prob}\ (\Lambda(P_{T,i}') < \Lambda(P_T)) \tag{6}$$

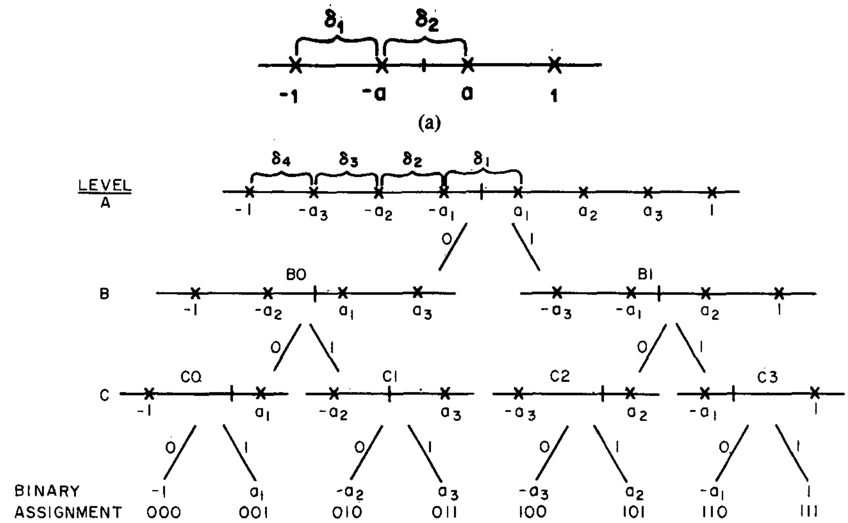
where $P_{T,i}'$, $i = 1, \cdots, M_T$ are all paths, which diverge from, and in less than or equal to T iterations later merge with, P_T. Denoting the set of all possible transmitted sequences as $\boldsymbol{P_T}$, the number of elements in $\boldsymbol{P_T}$ as $|\boldsymbol{P_T}|$, and the distance between the paths $P_{T,i}'$ and P_T as $D(P_T, P_{T,i}')$, as defined in (4), the probability of at least one error event is

$$P_E \le \lim_{T\to\infty} \frac{1}{|\boldsymbol{P_T}|} \sum_{P_T \in |\boldsymbol{P_T}|} \sum_{i=1}^{M_T} \text{Prob}\ (\Lambda(P_{T,i}') < \Lambda(P_T))$$

$$\le \lim_{T\to\infty} \sum_{i=1}^{M_T} \left[\sum_{P_T \in |\boldsymbol{P_T}|} \frac{1}{2|\boldsymbol{P_T}|} \exp\left(-\frac{D(P_{T,i}', P_T)}{8\sigma^2} \right) \right]. \tag{7}$$

The inner sum assumes a specific incorrect path and averages over all possible transmitted sequences of length T. This bound is straightforward to compute; however, it is computationally intensive.

Let $\boldsymbol{y}_i$ be the binary output vector of the coder ($m+1$ bits) at the ith interval, which is mapped to the channel symbol s_i, and suppose that path P_T corresponds to outputs $\{\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_T\}$, and $P_{T,i}'$ corresponds to outputs $\{\boldsymbol{y}_1 \oplus \boldsymbol{e}_1^i, \cdots, \boldsymbol{y}_T \oplus$



(b)

Fig. 2. (a) 4-AM constellation. (b) 8-AM constellation showing mapping of coded bits to levels by set-partitioning.

$e^i_T\}$, where e^i_k is an error vector of $m + 1$ bits, $1 \leq k \leq T$. It can easily be shown that the outer sum in (7) is over all possible error sequences $\{e_1, e_2, \cdots, e_T\}$, which correspond to diverging paths from a transmitted path of all zeros (i.e., the elements of the output vectors y are zero), and the inner sum is over all possible source sequences $\{b_1, b_2, \cdots, b_T\}$, which produce the outputs $\{y_1, \cdots, y_T\}$. Let $d(y, y')$ be the Euclidean distance between the channel symbols corresponding to y and y'. Then

$$D^2(P_T, P'_{T,i}) = \sum_{j=1}^{T} d^2(y_j, y_j \oplus e^i_j) \tag{8}$$

and (7) can be rewritten as

$$\begin{aligned}P_E &\leq \lim_{T\to\infty} \sum_{i=1}^{M_T} \left[\frac{1}{2|P_T|} \cdot \sum_{P_T \in P_T} \prod_{j=1}^{T} \exp\left(-\frac{d^2(y_j, y_j \oplus e^i_j)}{8\sigma^2}\right)\right] \\ &= \lim_{T\to\infty} \sum_{\{e_1, \cdots, e_T\}} \left[\frac{1}{2|P_T|} \cdot \sum_{\{y_1, \cdots, y_T\}} \prod_{j=1}^{T} \exp\left(-\frac{d^2(y_j, y_j \oplus e_j)}{8\sigma^2}\right)\right] \\ &= \lim_{T\to\infty} \frac{1}{2} \sum_{\{e_1, \cdots, e_T\}} \left[\prod_{j=1}^{T} \left(\frac{1}{2^m} \cdot \sum_{y_j} \exp\left(-\frac{d^2(y_j, y_j \oplus e_j)}{8\sigma^2}\right)\right)\right]\end{aligned} \tag{9}$$

where the factor 2^{-m} accounts for the number of equally likely source bit sequences, and the inner sum is over the 2^m possible coder outputs that can be generated at the jth interval. Because y_j in general can take on one of 2^{m+1} possible values, it is not immediately clear how the average over the output symbols y_j can be computed without enumerating all possible source sequences. (An *upper bound* for the last expression (9) is given in [8].)

Mapping coded output bits to constellation points by set-partitioning is described in [1] and [4]. Fig. 2 illustrates this procedure for the 8-AM constellation. If mapping by set-partitioning is assumed, then it is easily shown that the expression (9) is *equivalent* to

$$P_E \leq \lim_{T\to\infty} \frac{1}{2} \sum_{\{e_1, e_2, \cdots, e_T\}} \prod_{j=1}^{T} \left[\frac{1}{2^{m+1}} \cdot \sum_{y} \exp\left(-\frac{d^2(y, y \oplus e_j)}{8\sigma^2}\right)\right], \tag{10}$$

where the inner sum is over all possible (2^{m+1}) channel symbols, and does not depend on a particular transmitted path. This expression holds for some (but not all) nonuniformly spaced constellations (i.e., the constellations shown in Fig. 2), as well as for uniformly spaced higher dimensional constellations. The average over y can only take on 2^{m+1} values, corresponding to the different values of e_j. The sum is easily computed using the error state transition matrix [2].

III. Search for Optimal Codes

The search procedure used here is analogous to that used by Ungerboeck [1], except that here the union bound (10) is the optimality criterion, rather than minimum distance (i.e., the minimum Euclidean distance between two distinct paths in the trellis). Nevertheless, for a fixed number of trellis states it is assumed that a code which minimizes (10) is also optimal in the sense of having the maximum possible minimum distance. The search procedure therefore seeks those codes with maximum free distance and then evaluates the union bound at a specific SNR (17–18 dB, depending on the number of states in the trellis). As in [1], the search procedure is exhaustive with exclusion rules (although not all of the exclusion rules stated in [1] are applicable here). The heuristic code design rules in [1] reduce the number of codes in the search to approximately $2^{2(\nu-2)}$, where ν is the coder memory, and 2^ν is the number of states in the trellis.

A rate $m/(m+1)$ binary convolutional coder (i.e., m input bits, $m+1$ output bits) can be described by the input–output relation

$$y = Gx \tag{11}$$

where

$$y' = [y_1 y_2 \cdots y_{m+1}] \tag{12a}$$

is the vector of output bits at a particular time interval T,

$$x' = [x_1 x_2 \cdots x_{m+\nu}] \tag{12b}$$

is the vector of source bits in the shift register at time T, G is an $(m+1) \times (m+\nu)$ matrix that represents the location of shift register taps, and "prime" denotes transpose. The optimal codes have parallel transitions, or one uncoded bit, for $\nu \leq 7$. Consequently, G has dimension $2 \times (\nu + 1)$, corresponding to a rate 1/2 convolutional code. Table I shows the first and second rows of G, denoted as $g(0)$ and $g(1)$, respectively, in octal for the optimal codes found in the search. The 16-state code ($\nu = 4$) is shown in Fig. 3. Parallel transitions limit the squared minimum distance for codes with $\nu \geq 7$ to $16\delta^2$. Additional results [9] indicate that 2^{11} states are needed in order to provide a squared minimum distance of $17\delta^2$.

Fig. 4 shows the bound on error probability versus SNR for the codes shown in the table and for uncoded 4-AM transmission. The curves shown here were compared to the analogous curves for the codes found in [1]. The comparison indicates that for $\nu = 4$ and $\nu = 6$ the codes shown here offer a slight relative gain in SNR (approximately 0.2 dB). It should be mentioned, however, that the code listed in [1] for $\nu = 4$ has a coding gain of 4.2 dB, whereas the code shown here has a coding gain of 4.6 dB.[2] The curves in Fig. 4 for $\nu = 2, 3, 5$, and 7 were found to be nearly the same as the analogous curves for the codes listed in [1]. The table indicates that the savings in SNR due to coding at an event error probability of 10^{-6} is significantly less than the coding gain. This is partially due to the fact that the bound (5) used to derive (10) is pessimistic by approximately 0.5 dB at $P_E \approx 10^{-6}$. Nevertheless, the *relative* gain in SNR shown in the table for codes with different values of ν should be accurate.

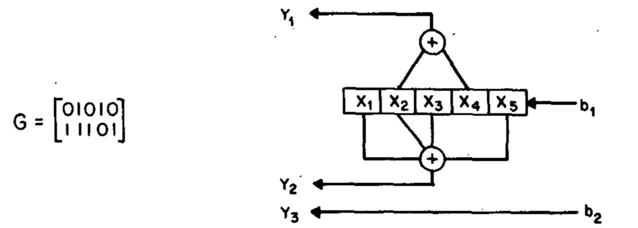
IV. Conclusions

An upper bound on the event error probability, rather than minimum distance, has been used as the optimality criterion for code searching. The codes found here offer only a slight savings in SNR compared to the codes found in [1] (less than 0.2 dB). A gradient search procedure has also been used to optimize the spacing between 4-AM and 8-AM signal levels assuming the specific four-state and 16-state codes shown in the table. The optimal value of a shown in Fig. 2(a) was found to vary with SNR and ν, ranging from just less than 0.4 to 0.8. The savings in SNR due to optimization of the signal levels was found to be about 0.2 dB for the four-state code and slightly less than 0.2 dB for the 16 state code at $P_E \approx 10^{-6}$. Because of the smaller spacings between points in the 8-AM constellation, the reduction in error probability that results from optimizing the signal levels is very small (less than 0.1 dB). The optimized 8-AM spacings are close to uniform in all cases. Optimization of signal levels with respect to minimum distance has also been studied in [10]. It is expected that the potential gains to be realized by optimizing other (two-dimensional) constellations are also minor.

[2] Also, the code listed in [1] for $\nu = 5$ has the same coding gain as the code shown here (4.9 dB), rather than 5 dB, as shown in [1].

TABLE I
8-AM Codes

ν	g(0)	g(1)	Coding Gain (dB)	SNR Gain ($P_E \approx 10^{-6}$)
2	2	7	3.31	2.1
3	4	13	3.77	2.5
4	12	35	4.56	2.8
5	10	55	4.91	3.1
6	54	161	5.23	3.5
7	152	323	5.81	3.8



MINIMUM DISTANCE SOURCE SEQUENCE: 1 0 1 1 0 0 0 0
EUCLIDEAN DISTANCE INCURRED AT EACH TRANSITION: 2δ δ 0 0 δ δ δ 2δ

Fig. 3. 16-state, 8-AM trellis code.

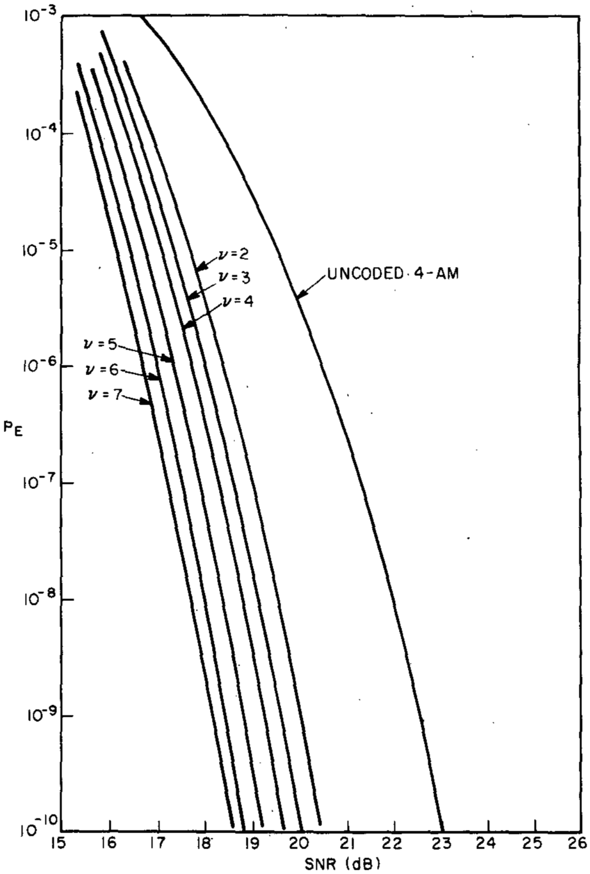


Fig. 4. P_E bound versus SNR for best 8-AM codes.

References

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

[2] A. S. Acampora, "Analysis of maximum likelihood sequence estimation performance for quadrature amplitude modulation," *Bell Syst. Tech. J.*, vol. 60, pp. 865–885, July–Aug. 1981.

[3] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. New York: McGraw-Hill, 1979.
[4] E. Biglieri, "High-level modulation and coding for nonlinear satellite channels," *IEEE Trans. Commun.*, vol. COM-32, pp. 616–626, May 1984.
[5] "A brief tutorial note on 64-point QAM with channel coding," AT&T Inform. Syst. proposal to CCITT, Jan. 1983.
[6] H. K. Thapar, "Real-time application of trellis coding to high-speed voiceband data transmission," *IEEE J. Select. Areas Commun.*, vol. SAC-2, pp. 648–658, Sept. 1984.
[7] M. L. Honig, "Performance analysis of multi-level trellis codes with additive Gaussian noise and intersymbol interference," submitted for publication.
[8] A. R. Calderbank and J. E. Mazo, "A new description of trellis codes," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 784–791, Dec. 1984.
[9] M. L. Honig, "Reduced complexity searching for multi-level trellis codes," Bell Commun. Res., Tech. Rep. 85-21241-51, 1986.
[10] D. Divsalar and M. K. Simon, "Combined trellis coding with asymmetric modulation," in *Proc. Globecom*, New Orleans, LA, Dec. 1985.

Correlation Properties of Integrated Random Variables

SHERMAN KARP

Abstract—**In general one does not measure point processes, but rather measures an integral of the process over a specific interval. In this note we have investigated how the covariance of a specific class of processes is altered by a finite measurement.**

In certain sensor systems it is common to compare outputs, each of which occurs at a different point but is averaged over the same interval. An example is the detection of a moving target. The moving target indication (MTI) is determined by taking the difference of each cell output at consecutive time intervals. The time intervals are adjusted so that the target moves one resolution cell during the integration. It is also assumed that the clutter background is slowly varying so that the only difference is due to the presence or absence of a target. Since there can never be perfect cancellation, there will always be a level of clutter leakage into the difference channel and the correlation of the noise in the two channels will be less than unity. In this note we will consider a class of covariance functions and compute the correlation of the two outputs as a function of time difference and time interval. Another example is the finite window opening during any sampling time.

Consider the intervals $(0, t_1)$, (t_1, t_2), $(t_2, t_2 + t_1)$ with $0 \leqslant t_1 \leqslant t_2$. Let the random variable x_t be defined on $(0, t_1)$ and the random variable y_t be defined on $(t_2, t_2 + t_1)$. Let

$$x_{t_1} = \frac{1}{t_1} \int_0^{t_1} x_t \, dt$$

$$y_{t_2} = \frac{1}{t_1} \int_{t_2}^{t_2+t_1} y_t \, dt$$

and form

$$E[x_{t_1} \; y_{t_2}] = E\left[\frac{1}{t_1^2} \int_0^{t_1} \int_{t_2}^{t_2+t_1} x_t y_{t'} \, dt \, dt'\right.$$

$$\left. = \frac{1}{t_1^2} \int_0^{t_1} \int_{t_2}^{t_2+t_1} E[x_t y_{t'}] \, dt \, dt' \right]$$

$$= \rho(t_1, t_2).$$

Consider the covariance function

$$E[x_t y_{t'}] = e^{-\alpha|t'-t|} \cos \beta(t'-t).$$

For this case we have

$$\rho(t_1, t_2) = e^{-\alpha t_2} \operatorname{Re}\left[e^{j\beta t_2} \left\{ \frac{\sinh\left[\frac{(\alpha - j\beta)}{2} t_1\right]}{\frac{(\alpha - j\beta)}{2} t_1} \right\}^2 \right].$$

Notice that as $t_1 \to 0$ the two variables decorrelate as

$$e^{-\alpha t_2} \cos \beta t_2$$

as expected. Also notice that $\beta = 0$ corresponds to the covariance function of a first-order Markov process, for which

$$\rho(t_1, t_2) = e^{-\alpha t_2} \left[\frac{\sinh \frac{\alpha t_1}{2}}{\frac{\alpha t_1}{2}} \right]^2$$

and falls more rapidly than $e^{-\alpha t_2}$, as one might expect. Similarly for $\alpha = 0$

$$\rho(t_1, t_2) = \cos \beta t_2 \left(\frac{\sin \frac{\beta t_1}{2}}{\frac{\beta t_1}{2}} \right)^2$$

and we have a windowing effect due to the integration time t_1.

One might also consider exploiting different properties of the target process with two sensors, each observing one of the properties. In this case it is possible to have overlapping time intervals where now $0 \leq t_2 \leq t_1$ and we have $(0, t_2)$, (t_2, t_1),

Paper approved by the Editor for Image Processing and Pattern Recognition of the IEEE Communications Society. Manuscript received July 16, 1985.
The author is with the Strategic Technology Office, Defense Advanced Research Projects Agency (DARPA), Arlington, VA 22209.
IEEE Log Number 8609573.