PERFORMANCE OF BINARY PSK COMMUNICATION SYSTEMS[1]

J. F. OBERST and D. L. SCHILLING
Department of Electrical Engineering
Polytechnic Institute of Brooklyn.

Summary The degree of RF coherence which can be established between transmitter and receiver greatly influences the performance of binary communication systems. Practical systems are partially coherent; the main classes are transmitted reference (TR) and single channel (SC). Although SC systems are potentially superior, they are difficult to analyze and have an inherent mark-space ambiguity problem.

In this paper, four SC PSK systems have been studied using Monte Carlo simulation on an IBM 360/50 digital computer. Differential data encoding was used. The systems investigated include Decision Feedback (DF), Squaring (SQ), and a variation of SQ called Absolute Value (AB). In addition, a new Maximum Likelihood (ML) SC system, which is optimum in a restricted sense, is derived and simulated.

The simulation results show that all of these systems yield comparable average probability of error. This is in contrast with results which have been published previously. Furthermore, the systems can all be shown to reduce to Differential PSK when the number of reference bauds is one. Finally, a method is introduced for studying the effects of various methods of data encoding on SC system operation.

Partially Coherent PSK Systems

The degree of radio frequency (RF) coherence which can be established between transmitter and receiver of a communication system greatly influences the performance attainable. Assuming that the signal amplitude (given i) and the baud timing, or bit sync, is known at the receiver, we can write the RF signal received in the baud [0, T] as

$$v(t) = s_i(t, \theta) + n(t) \qquad t \in [0, T]$$

where the subscript i indicates that $s_i(t, \theta)$ is the signal corresponding to message m_i, i = 1, 2. n(t) is white gaussian noise with one-sided power spectral density N_o. θ is the RF

[1] This work was supported by the National Aeronautics and Space Administration under Grant NGR-33-006-020, and forms part of a Ph.D. (E.E.) Dissertation [1] submitted to the Polytechnic Institute of Brooklyn.

phase, assumed to be constant over the baud, which may be known with varying degrees of accuracy depending on the particular communication system. The accuracy with which the value of θ is known at the receiver determines the degree of RF coherence and thereby the minimum average probability of error P(E) which can be achieved.

Two important signal parameters are the energies of the two possible received signals E_i, and their mutual correlation coefficient ρ:

$$\rho = (E_1 E_2)^{-1/2} \int_0^T s_1(t, \theta)\, s_2(t, \theta)\, dt .$$

If the actual value of θ is known accurately at the receiver, the system is termed coherent. It is known that the best signal set for such a system, in the sense of minimum P(E), consists of a pair of equal energy anticorrelated (ρ = -1) signals. Phase shift keyed sinusoids, defined by

$$s_1(t, \theta) = C \cos(\omega_o t + \theta)$$
$$s_2(t, \theta) = C \cos(\omega_o t + \theta + \pi) = - C \cos(\omega_o t + \theta) ,$$

form such a set. With the signal energy E received in each baud and the signal-to-noise ratio R defined as

$$E = \frac{C^2 T}{2} \qquad\qquad R = \frac{E}{N_o}$$

it can be shown that the minimum P(E) attainable is

$$P(E) = \frac{1}{2} \operatorname{erfc} \sqrt{R} \tag{1}$$

The receiver which attains this P(E) is the correlation detector shown in Fig. 1.

In practical communication systems the received phase is not known with perfect accuracy. Often, however, an estimate $\hat{\theta}$ of θ is available, and systems using such an estimate are called partially coherent. When the estimate is good, and the signals are properly chosen, partially coherent systems may attain a P(E) close to the lower bound (1). Unfortunately, the problem of overall system design, including signal specification and receiver design, is quite complex. The difficulty is that the phase estimate is usually derived at the receiver by special processing of the signals received over several bauds. Thus considerations of receiver complexity, signal structure, and channel characteristics all interact. The procedure which has led to many useful partially coherent systems is to postulate a “reasonable” signal set on the basis of the method to be used by the receiver to derive the phase estimate, and then to seek the parameter values which yield the best operation. Since anticorrelated signals are generally superior when a good phase estimate is available, the best choices are usually modifications of the simple coherent PSK system discussed above. These systems usually fall into one of two categories, transmitted reference and single channel, although combinations of the two are also possible.

Single Channel Systems

Transmitted Reference (TR) systems transmit a reference signal to the receiver through a channel separate from the data channel. Thus the transmitted signal is a combination of a normal PSK signal and a signal to be used only for RF phase synchronization. The major disadvantage of TR systems is that the available power at the transmitter must be shared between these two signals. In single channel (SC) systems a normal, fully modulated PSK signal is transmitted. The phase estimation portion of the receiver processes this PSK signal in some nonlinear manner in order to remove the modulation. The various systems differ in the nature of this nonlinear processing. The great advantage of SC systems is that the entire transmitter power is used for both data transmission and phase synchronization. Thus SC systems are potentially superior to TR systems.

It is appropriate at this point to introduce the mathematical model used in this study of SC systems and to note the two assumptions involved. First, the carrier frequency ω_o is assumed to be known accurately at the receiver. Second, the quality of the phase estimate obtainable in actual systems is limited by the rate of change of received phase, caused by channel variations, oscillator phase jitter, etc. This limitation is incorporated into a simple model by postulating a constant received phase, but a limited allowable measurement time. This measurement time is taken as qT, extending over q bauds of duration T. The waveform available at the receiver for the decision on the data transmitted for $t \in [0, T]$ may therefore be represented as

$$v(t) = m(t)\, C \cos(\omega_o t + \theta) + n(t), \qquad t \in [-qT, T]. \tag{2}$$

where m(t) represents the modulation, and takes on the values ± 1 in each baud independently and with equal probability. Some receivers also store the last q decisions.

Although the model (2) is simple, two major difficulties are encountered in the study of SC systems. The first involves the nonlinear processing of the signal performed by the receiver in order to construct a partially coherent reference waveform. Such processing makes it extremely difficult to obtain results analytically. For this reason, Monte Carlo simulation on a digital computer is useful in studying system performance. The second difficulty encountered in SC systems is the “mark-space ambiguity” problem. In a simple SC system there is no way of determining the absolute sense of the signal received in each baud. That is, even though in low noise the bauds during which a 1 was transmitted can be distinguished from those during which a 0 was transmitted, it is impossible to determine which of these two groups of bauds actually corresponds to the 1 signal. In most SC PSK receivers, this mark-space ambiguity takes the form of two fairly stable average phase values of the derived reference, separated from each other by π radians. Even if the correct reference sense is established at the beginning of the transmission, noise may cause loss of synchronism and eventual reference reversal during the transmission, inverting the decoded message. Several techniques are available for dealing with this problem, the simplest of which is differential data encoding. As shown in Fig. 2, data of 1 is represented by no phase change from the signal transmitted in the previous baud, and data of 0 by a phase change of π radians. Since the presence or absence of a phase transition between bauds can be detected equally well using a reference with phase near θ or θ + π, there can be no inversion of the data with this technique. However, in a low error rate system where adjacent decisions are approximately independent, differential encoding almost doubles P(E).

Simulation Technique

There are three basic reasons that Monte Carlo simulation on a digital computer is useful for studying SC PSK systems. First, as mentioned previously, SC systems are difficult to analyze exactly because the received signal is processed in a complicated way. Second, the systems studied are basically discrete, and are therefore easy to model on a digital computer. Furthermore, these systems process independent, normal random variables which are easy to generate by a computer subroutine. Finally, sophisticated error counting and analysis can be included in the simulation programs without much additional effort.

The normally distributed random numbers used in the simulations were obtained from uniformly distributed random numbers by a mathematical transformation of probability density functions. The uniform random numbers were generated by the multiplicative method, using the relationship

$$y_i = a\, y_{i-1} \pmod{m}. \qquad (3)$$

For the IBM 360/50 used, the parameters in (3) which gave good results are

$$m = 2^{31} \qquad a = 2^{17} + 3 \qquad y_o = 16779715\,.$$

In all of the systems simulated, the random variables at the receiver input which must be operated upon are independent and normal, with p. d. f.

$$\begin{aligned} w_1 &= N(m_i\sqrt{2R}\ \cos\theta,\ 1) \\ w_2 &= N(m_i\sqrt{2R}\ \sin\theta,\ 1)\ . \end{aligned} \qquad (4)$$

$m_i = \pm 1$ represents the modulation. Since the performance of the receivers studied is identical for all possible messages and for any received phase, (4) can be simplified by assuming

$$m_i = +1 \qquad\qquad \theta = 0 .$$

The random numbers needed for the simulation become

$$w_1 = N(\sqrt{2R}\ ,\ 1)$$
$$w_2 = N(0,\ 1) .$$

The simulated systems were started with noiseless reference vectors. Since a large number of trials (typically 10,000) are performed for each value of R, the choice of initial reference vector should not influence the error count appreciably. Finally, it should be noted that the various Systems simulated had identical random numbers as inputs, so the results are truly comparable.

Results for Various Systems

Differential data encoding was used in all of the simulations. The systems studied include Differential (D) PSK, Decision Feedback (DF) PSK, Squaring (SQ) PSK, and a variation of SQ-PSK called Absolute Value (AB) PSK. In addition, a new Maximum Likelihood (ML) SC system, which is optimum in a restricted sense, is derived and simulated.

We first establish a performance bound for differentially encoded SC PSK systems as follows. A SC system is operating ideally when its derived reference is noiseless. Then the probability p of incorrectly identifying the sign of the signal received in each baud is given b (1):

$$p = \frac{1}{2}\ \mathrm{erfc}\sqrt{R} .$$

Since the reference is perfect, successive decisions are independent, and since differential encoding is used, an error occurs when a wrong detection is followed by a correct detection or vice versa. Therefore the minimum average error probability is given by

$$P(E) = p(1 - p) + (1 - p)p = \mathrm{erfc}\sqrt{R}\,(1 - \frac{1}{2}\ \mathrm{erfc}\sqrt{R}) . \qquad (5)$$

This bound is useful in evaluating the simulation results.

We now consider Differential (D) PSK which is the simplest and in certain ways the basic SC PSK communication system. In D-PSK, the signal received during the previous baud serves as the partially coherent reference for the present decision. A D-PSK receiver is shown in Fig. 3. Analytical results are available [1] for various error probabilities in DPSK:

$$P(E) = \frac{1}{2} e^{-R}$$

$$P(E_{j+1}/E_j) = \frac{1}{2\pi} \int_0^{\pi} \mathrm{erfc}^2 (\sqrt{R} \cos\phi) \left\{ 1 + \sqrt{R\pi} \cos\phi \, e^{R \cos^2 \phi} [1 + \mathrm{erf}(R \cos\phi)] \right\} d\phi$$

$$P(E_{j+2}/E_j) = \frac{1}{2} e^{-R} \qquad (6)$$

The results of the error counts from the simulation are shown in Figs. 4 and 5, and the analytical curves (6) are included for comparison. The excellent agreement in these two figures indicates that the random number generator used produces numbers with the desired distribution and independence.

In Decision Feedback (DF) PSK, combination with the reference of the signal received during the present baud is delayed until a decision about which signal was transmitted i's made by the receiver. This decision is then used as if it were correct to modify the present signal before it is combined with the reference. Since usually P(E) is small, the rge majority of received signals are combined in the correct way. The structure of one DF-PSK receiver is shown in Fig. 6. The delay-and-multiply operation at the output is used to recover the data, which has been differentially encoded. It is easy to show that - when q = 1, the DF-PSK receiver of Fig. 6 reduces to the D-PSK receiver of Fig. 3. The error counts from the DF-PSK simulation are shown for three values of q in Fig. 7. It is clear that for medium values of q the system is operating close to its theoretical limit.

Now we examine a class of SC PSK systems known as Harmonic Tracking (HT). The general HT-PSK receiver is shown in Fig. 8. The operation of the reference deriving section is best explained by representing the random variables (r. v.) w_{1i} , w_{2i} in the ith baud by the vector $\underline{w}_i = w_{1i} + j\, w_{2i}$ in the complex plane. If ϕ_i is the phase error due to noise in the ith baud, the received phase is given by

$$\psi_i = \mathrm{Arg}\,(\underline{w}_i) = \tan^{-1} \frac{w_{2i}}{w_{1i}}$$

$$= \theta + \phi_i \quad \text{or} \quad \theta + \phi_i + \pi$$

depending on whether a 1 or 0 was transmitted. In either case, the phase doubler output $\underline{w}_i'$ has angle

$$\theta_i' = \mathrm{Arg}\,(\underline{w}_i') = 2\,\psi_i = 2\,(\theta + \phi_i)$$

which is independent of the data transmitted during baud i. The vectors $\underline{w}_i'$ corresponding to the q reference bauds are combined to give a resultant vector $\underline{w}'$ with phase

$$\theta' = \text{Arg}\,(\underline{w}') = 2\,(\theta + \phi')$$

where ϕ' represents the total error due to the noise. θ' is then halve to give the final phase estimate $\hat{\theta}$

$$\hat{\theta} = \text{Arg}\,(\underline{w}) = \theta + \phi' .$$

Two details of this technique demand further attention. First, the method of combining the $\underline{w}_i'$ has not been completely specified. For simplicity, we assume they are combined by vector addition. Then it remains to specify the exact action of the phase doubler. If $\underline{w}_i'$ is written in polar form,

$$\underline{w}_i' = M_i' \, e^{2j\psi_i} ,$$

it is clear that to specify the phase doubler we need only specify its effect on the amplitude of the input vectors

$$\underline{w}_i = M_i \, e^{j\psi_i} .$$

If the phase doubler consists of a vector squarer, we have

$$M_i' = M_i^2$$

which corresponds to SQ-PSK. In general, if the amplitude is raised to the ν th power, we have ν th law HT-PSK.

The second and most important consideration is that the procedure of halving the phase θ' is analogous to taking the square root of a complex number, and therefore the final phase estimate $\hat{\theta}$ has two possible values separated by π radians. A rule must be established for choosing one of these two values. A satisfactory rule is to choose the value of $\hat{\theta}$ which is closest to the $\hat{\theta}$ derived for the previous decision. This choice gives good results for all values of q, and can be shown to be optimum for q = 1, since HT-PSK is equivalent to D-PSK under these conditions.

The HT-PSK system described above was simulated on a digital computer for two values of ν . The error counts for $\nu = 2$ are shown in Fig. 9. This corresponds to the most widely studied system, SQ-PSK. The results for AB-PSK ($\nu = 1$) are shown in Fig. 10., The performance of both HT systems is nearly identical. When these results are compared with those for DF-PSK in Fig. 7, very little difference is discernible. DF seems to have a slight advantage for q = 2, but the systems all look equivalent for larger values of q. Proakis et. al. [2], on the other hand, have found on the basis of a similar digital simulation that DF-PSK is consistently superior to SQ-PSK. Apparently they did not simulate the best possible SQ system. For q = 1, they obtained an experimental P(E) for SQ-PSK significantly higher than the known result (6) for D-PSK, while we have shown that with proper design these two systems are identical.

The final system investigated is called Maximum Likelihood (ML) PSK. This system was derived in the following way. Assume that the differentially encoded PSK signal (Z) is available, but past decisions made by the receiver are not remembered. ML-PSK is the receiver which attains the smallest possible P(E) for the decision about the latest possible (at t = 0) phase transition. This is accomplished by applying a maximum a posteriori decision rule to the received signal. The ML receiver is quite complicated, its complexity growing exponentially with q, but it can be implemented. This complexity is reflected in the large amount of computer time used in the simulation. For this reason only two values of q were used, and fewer trials were made for this system than for the others. The error counts are shown in Fig. 11. Even for the relatively small q = 5, the experimental P(E) is nearly equal to the theoretical lower bound. Performance is consistently superior to the other SC systems considered, but is not good enough to justify the receiver complexity.

In summary, the simulation results show that the popular SC PSK systems perform similarly. The theoretical lower bound in P(E) for a differentially encoded SC PSK system is approached quite closely for moderate values of q. Furthermore, all of the systems studied reduce to Differential PSK for the special case q = 1.

Data Encoding for SC Systems

It is clear from the results of the simulations that differential data encoding nearly doubles P(E) over that attainable by an ideal coherent PSK communication system with direct data encoding. Therefore it is desirable to investigate other methods of combatting mark-space ambiguity in SC systems. The difficulty here lies in the complexity of the receivers being studied. Fig. 12 shows a simple model which approximates the behavior of most SC systems for reasonable values of q. System operation with normal and inverted phase reference is represented by states N and I, respectively. The probability of reference reversal in any baud is λ. The probability of a detection error in state N is p, and in state I is $\overline{p} = 1 - p$. Such errors are assumed to occur independently from baud to baud.

If $q \ll 1/\lambda$, which is the case for reasonable values of q and λ, the transitions between states of an actual system may be expected to occur randomly. Furthermore, for a first order Markov source like the proposed model, the number of bauds spent in each state are exponentially distributed. This was verified experimentally for DF-PSK. Also, if errors in each state occur independently, it is easy to verify that when differential encoding is used

$$P(E_{j+1}/E_j) = \frac{1}{2} .$$

This was found to be true for DF-PSK with $q \geq 2$ by simulation, as shown in Fig. 13. The parameters p and λ remain to be identified. Since this state model is proposed for comparing various methods of encoding, rather than for accurate calculation of P(E), it is sufficient to obtain only rough estimates of p and λ. An approximate result for p is available from (1):

$$p \cong \frac{1}{2} \operatorname{erfc} \sqrt{R} .$$

λ is a strong function of both R and q, and is usually best determined experimentally.

Space does not permit a comparison of the many possible data encoding schemes here. Some methods which have been studied include differential encoding, transmission of "check bits" singly and in groups, use of a decision feedback channel with delay from receiver to transmitter, and the use of specially selected error correcting codes. As an example of the technique employed, let us briefly consider the familiar scheme of differential encoding. An error occurs in this system when a correct decision is followed by an incorrect decision or vice versa. Therefore the avera e robabilit of decision error is

$$P(E) = (1-p)[\bar{\lambda} p + \lambda(1-p)] + p[\bar{\lambda}(1-p) + \lambda p] = \lambda + 2p(1-p)(1-2\lambda) .$$

When $\lambda << p$, this reduces to the result (5) obtained previously as a lower bound on system performance. The performance of other encoding schemes in terms of λ and p can be computed in a similar manner. Note that many of these schemes, such as the transmission of reference check bits, are actually combinations of single channel (SC) and transmitted reference (TR) systems. It is easy to show that under some conditions these hybrid systems achieve lower P(E) than is obtained with the usual SC technique of differential encoding.

References

1. Oberst, J. F., "Binary PSK Communication Systems," Ph. D. (E. E.) Dissertation, Polytechnic Institute of Brooklyn, September 1968.
2. Proakis, J. G., P. R. Drouilhet, Jr., and R. Price, "Performance of coherent detection systems using decision-directed channel measurement," IEEE Transactions on Communication Systems, March 1964, pp. 54-63.

3. Oberst, J. F., and D. L. Schilling, "Double error probability in differential PSK," Proceedings of the IEEE, June 1968, pp. 1099, 1100.

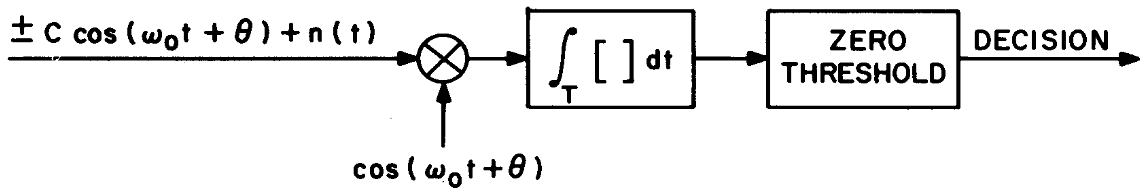

Fig. 1 - Correlation detector

MESSAGE SEQUENCE		0	1	1	0	1	1	0
DIRECT ENCODING		π	0	0	π	0	0	π
DIFFERENTIAL ENCODING	0	π	π	π	0	0	0	π

Fig. 2 - Direct and differential encoding

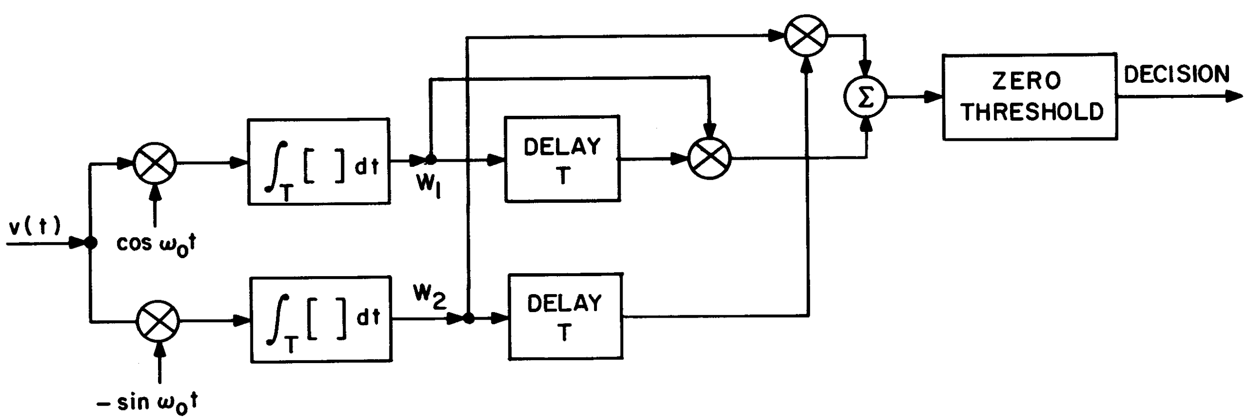

Fig. 3 - D-PSK receiver

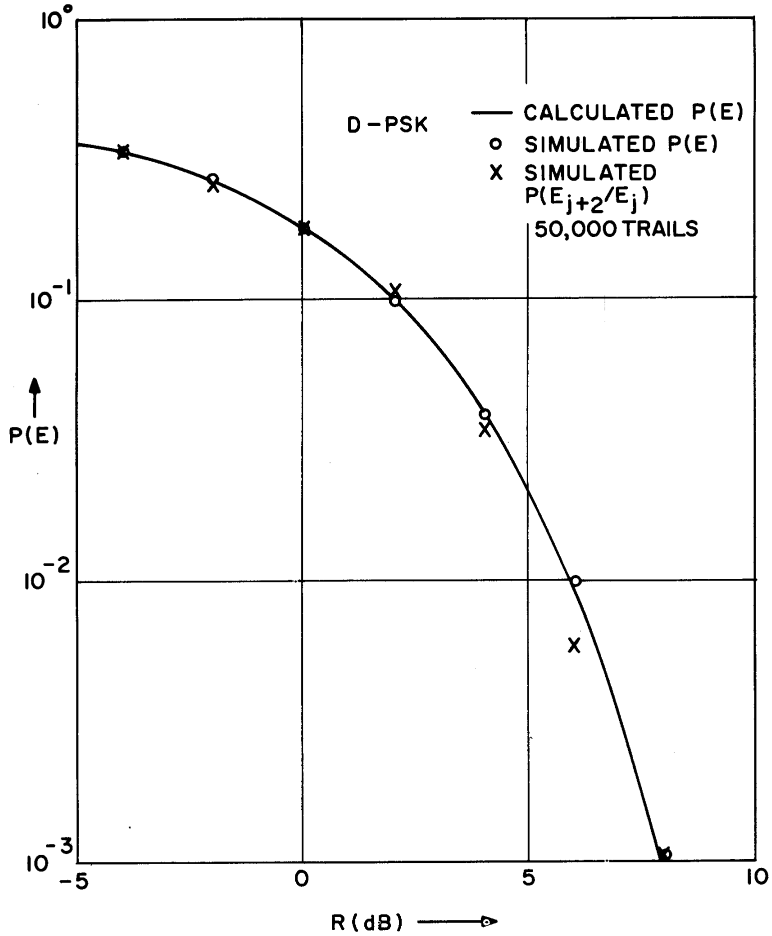

Fig. 4 - P(E) and $P(E_{j+2},/E_j)$ versus R for D-PSK

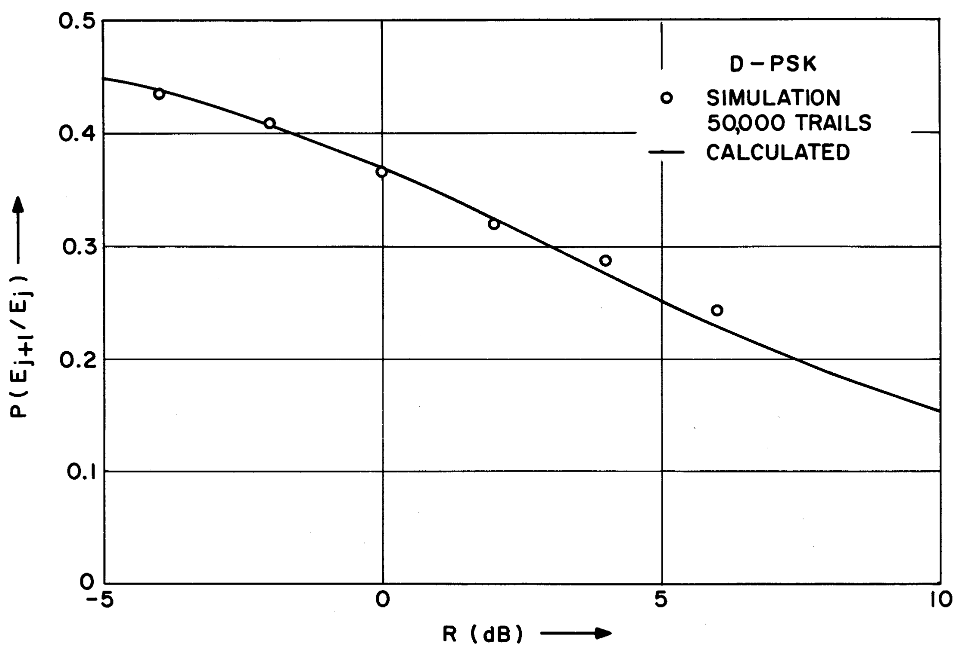

Fig. 5 - $P(E_{j+1}/E_j)$ versus R for D-PSK

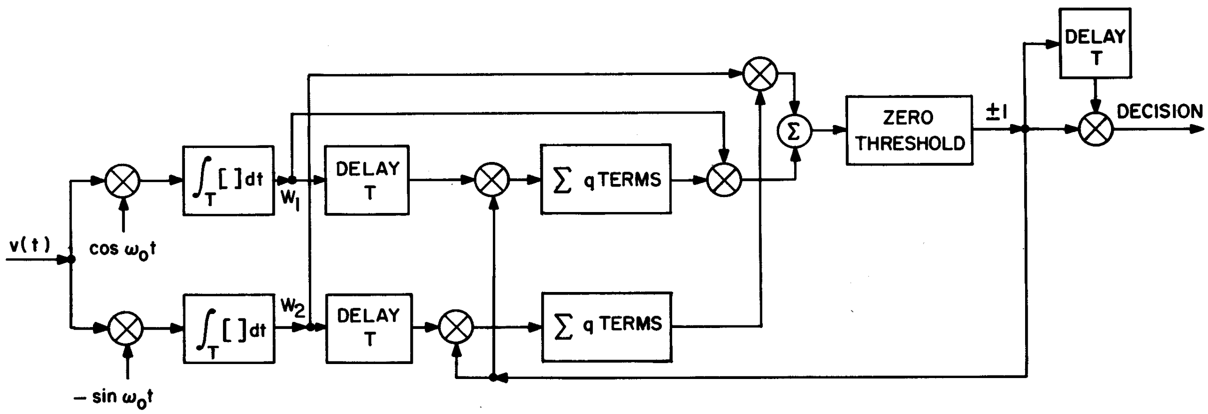

Fig. 6 - DF-PSK receiver

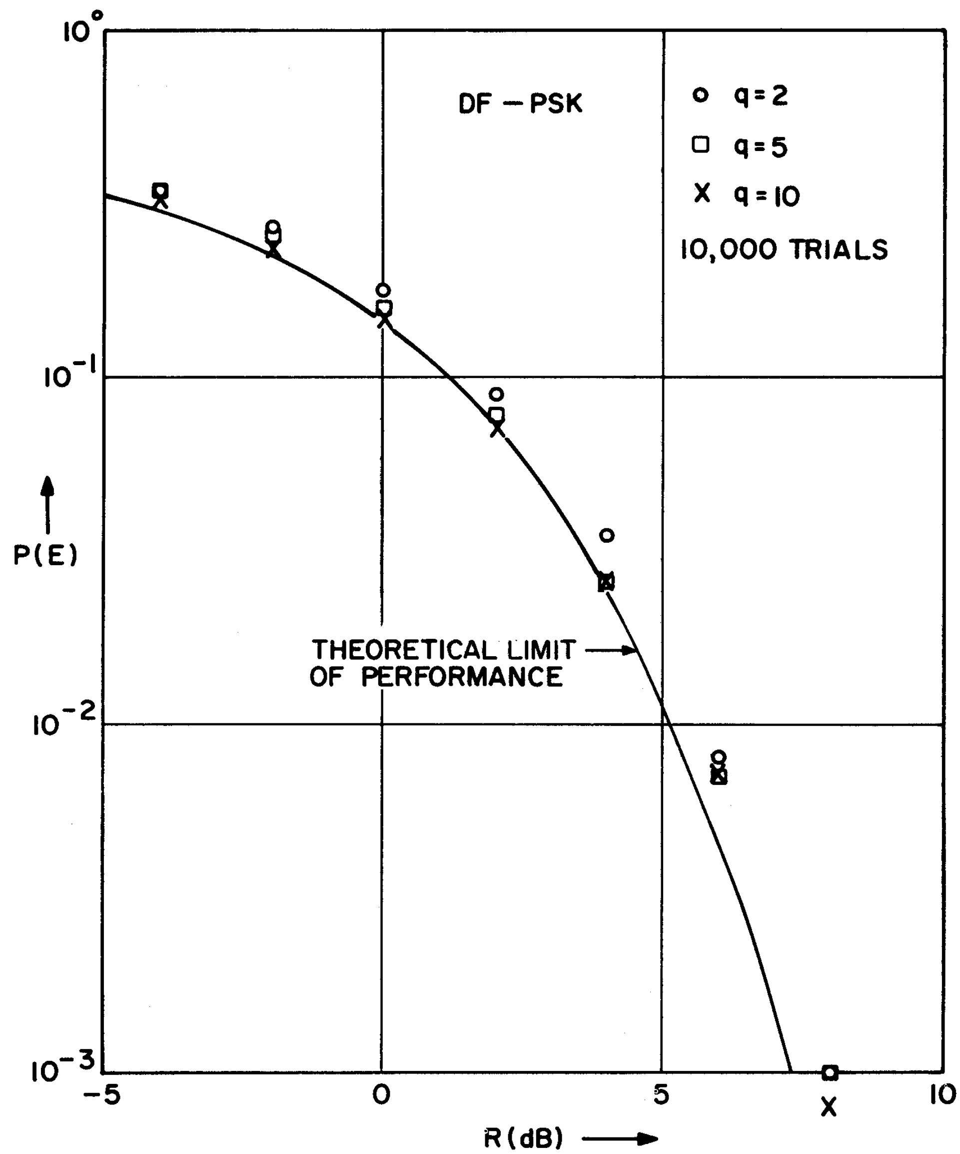

Fig. 7 - Simulation results for DF-PSK

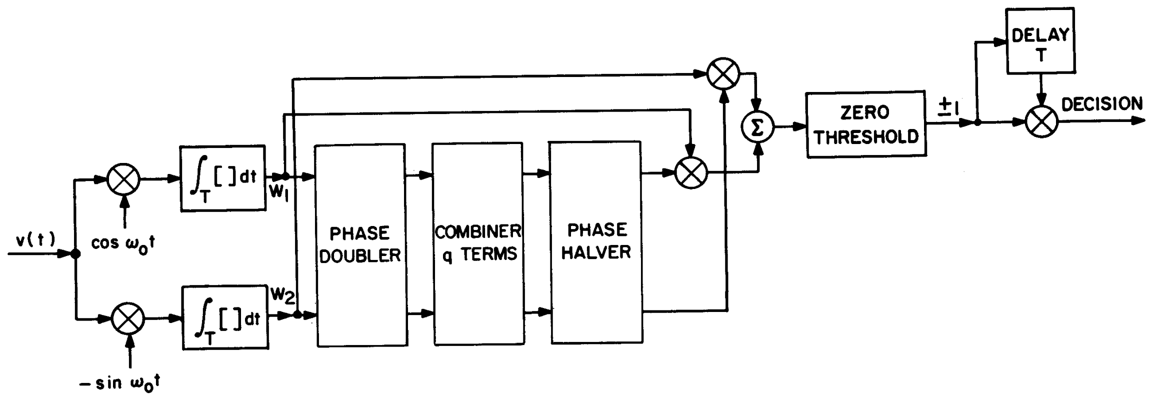

Fig. 8 - General HT-PSK receiver

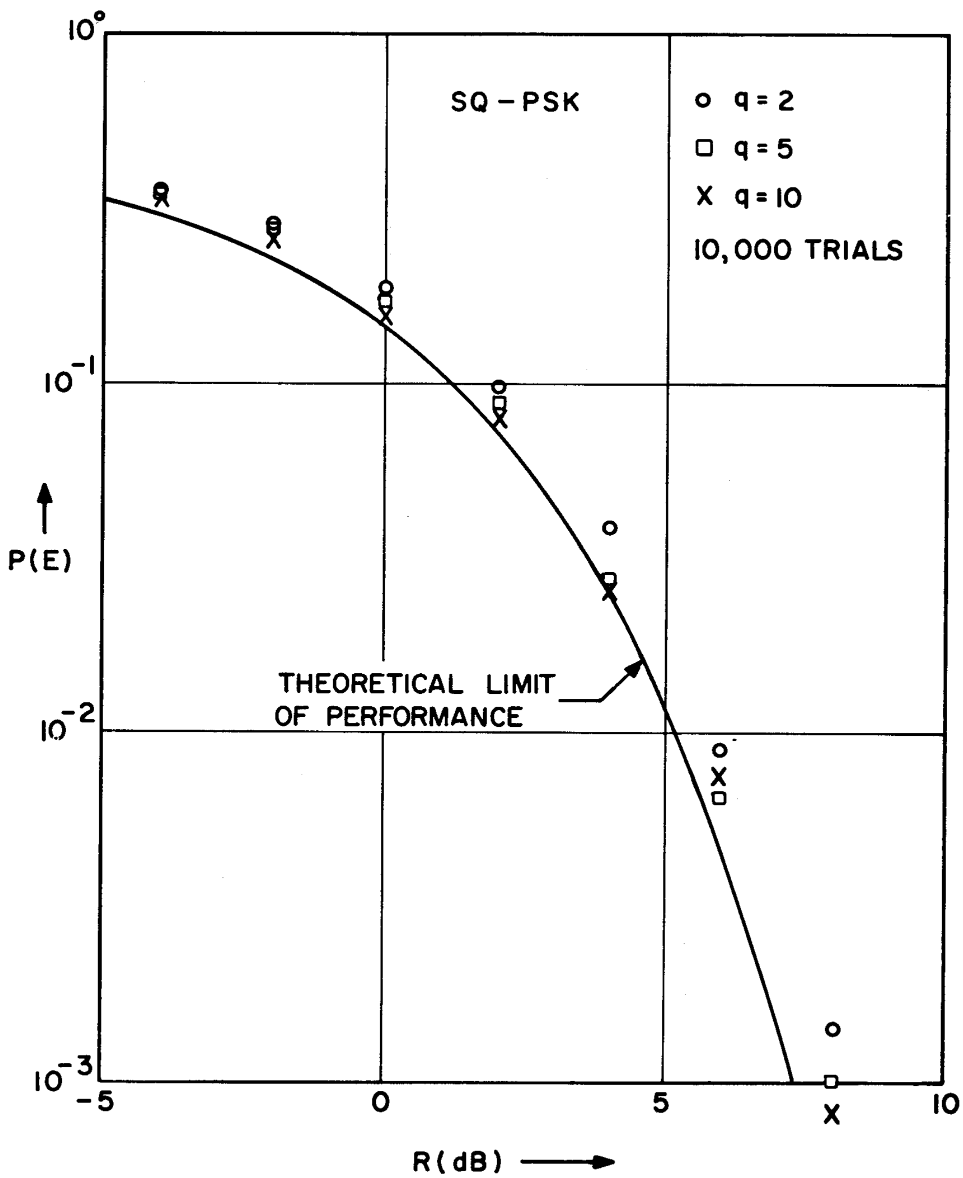

Fig. 9 - Simulation results for SQ-PSK

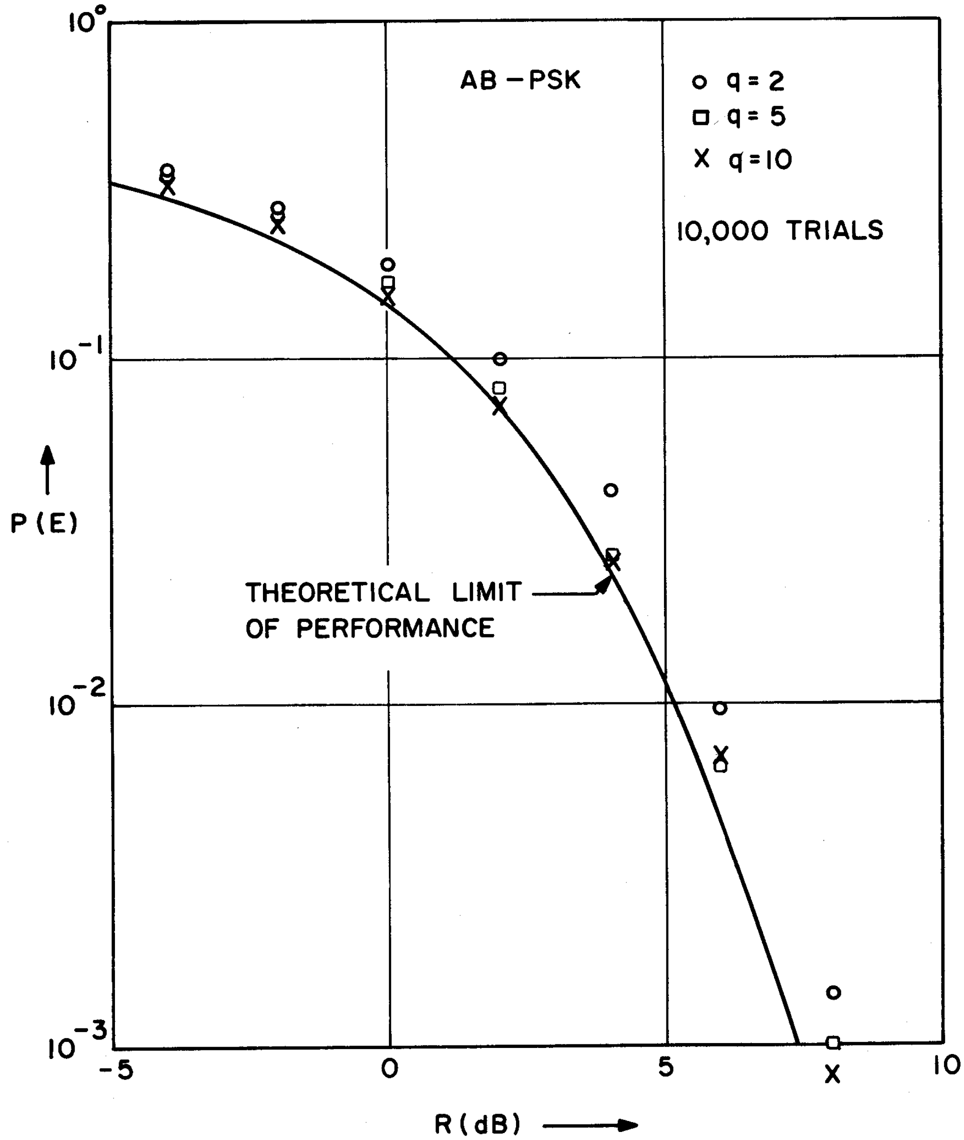

Fig. 10 - Simulation results for AB-PSK

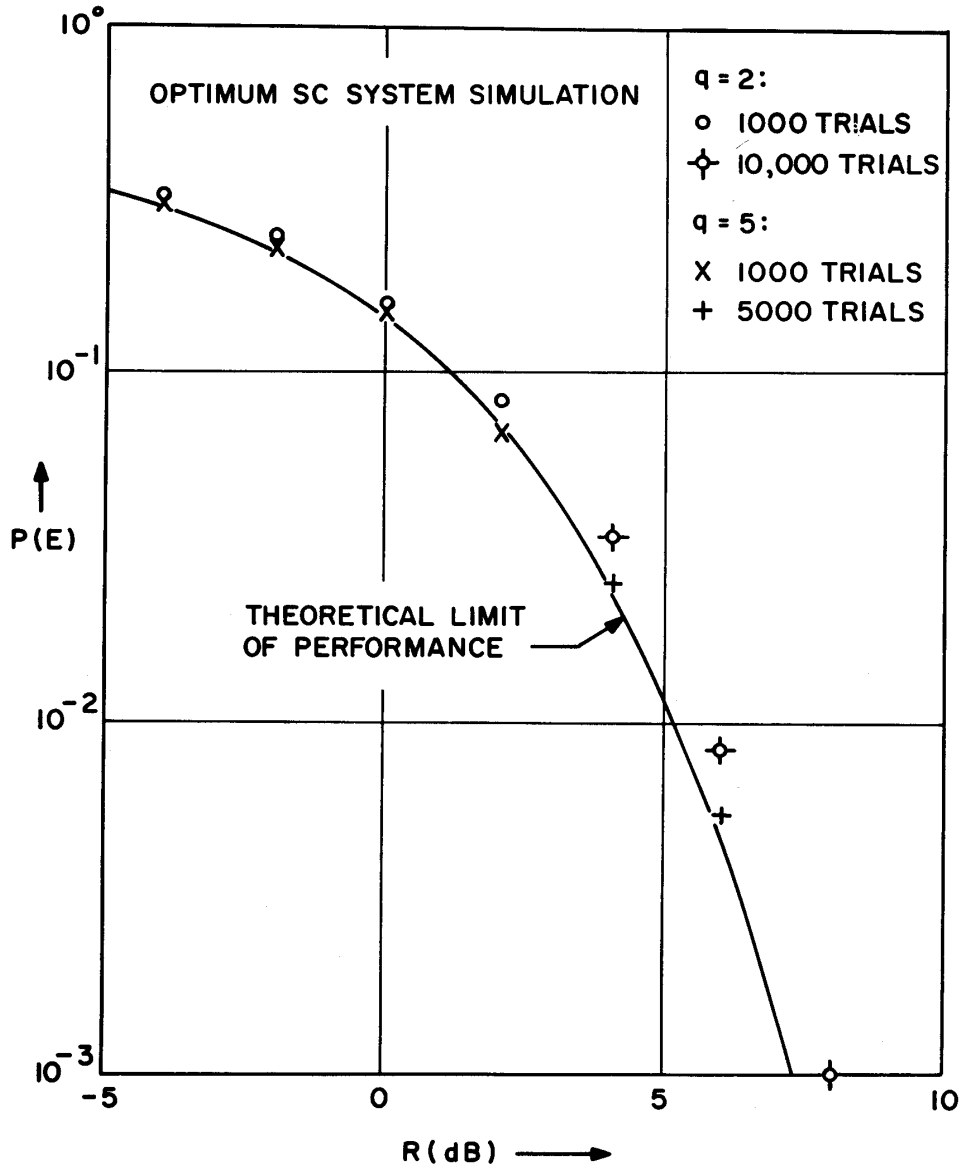

Fig. 11 - Simulation results for ML-PSK

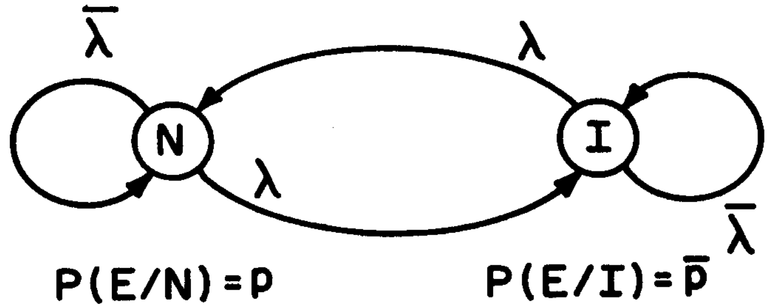

Fig. 12 - State model for SC PSK system

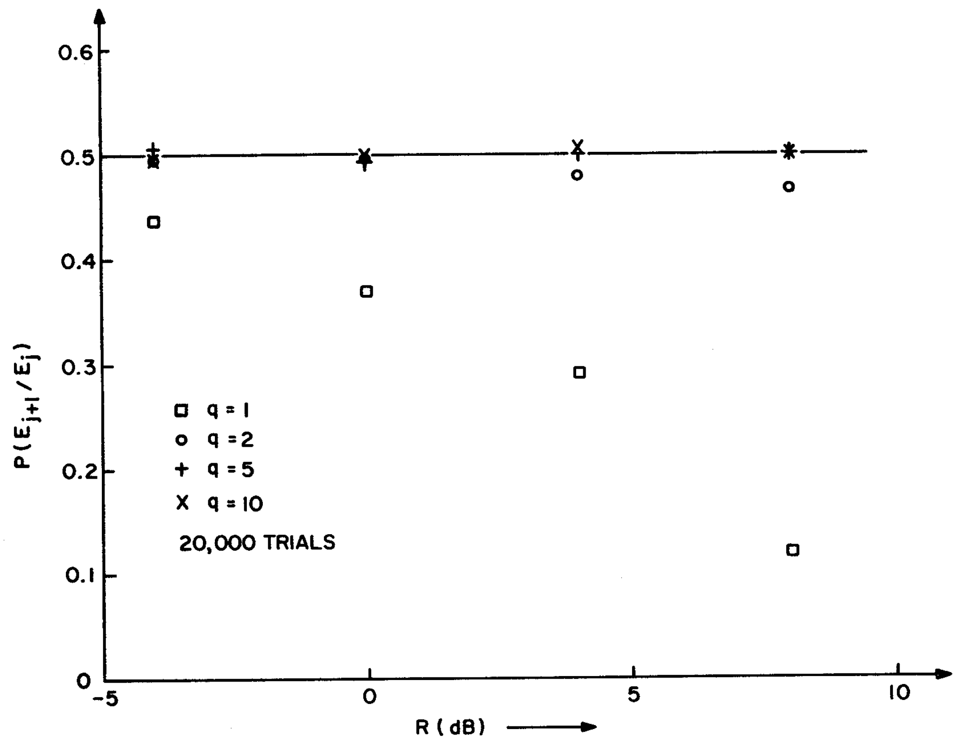

Fig. 13 - $P(E_{j+1}/E_j)$ for DF-PSK with differential encoding