8-1994

Optimum symbol-by-symbol detection of uncoded digital data over the Gaussian channel with unknown carrier phase

Pooi Yuen KAM

Seng Siew NG

Tock Soon NG
Singapore Management University, tsng@smu.edu.sg

DOI: https://doi.org/10.1109/26.310614

Optimum Symbol-By-Symbol Detection of Uncoded Digital Data Over the Gaussian Channel with Unknown Carrier Phase

Pooi Yuen Kam, *Senior Member, IEEE,* Seng Siew Ng, and Tok Soon Ng

***Abstract*—A theory of optimum receiver design for symbol-by-symbol detection of an uncoded digital data sequence received over the Gaussian channel with unknown carrier phase is presented. Linear suppressed-carrier modulation is assumed. The work here aims at laying a conceptual foundation for optimum symbol-by-symbol detection, and rectifies existing approaches to the problem. The optimum receiver structure is obtained explicitly for an arbitrary carrier phase model, but its computational requirements are too heavy in general for any practical implementation. In one important special case, namely, the case in which the carrier phase can be treated as a constant over some $K + 1$ symbol intervals, the optimum receiver can be approximated by a readily implementable decision-feedback structure at high SNR. Simulated error performance results are presented for this latter receiver for PSK modulations with various carrier phase models. Since a decision-feedback receiver can encounter a "runaway," a variation of this receiver is developed which uses feedforward of tentative decisions concerning future symbols. This modified receiver does not have any "runaway" problem, and has been shown to yield good error performance via simulations.**

I. Introduction

CONSIDER a sequence of uncoded digital data transmitted using linear suppressed-carrier modulations over an additive white Gaussian noise (AWGN) channel which, in addition, introduces an unknown carrier phase shift θ. For an uncoded sequence, each symbol in the sequence will be detected individually, i.e., symbol-by-symbol (SBS) with minimum symbol error probability (SEP). The decision on each symbol, however, will be based on the totality of received signals due to the transmission of the entire data sequence because the continuity of the carrier phase process introduces memory into the received signals. This paper serves to elucidate the explicit structure of this optimum SBS receiver, and shows systematically the signal processing required in making an optimum decision on each symbol of the sequence.

Kobayashi [1] is the first author to propose an unstructured approach to the problem of optimum data detection in the presence of an unknown carrier phase; unstructured meaning no assumptions are made a priori concerning the receiver structure. (He also includes unknown timing phase and intersymbol interference). Prior to [1], all data receivers for the unknown phase channel were designed using the structured approach. In this structured approach, the receiver is assumed to consist of a carrier phase estimator followed by a coherent detector (see Fig. 1). The carrier phase is first recovered by using some phase estimating system such as the Costas' loop. See, for instance, [2]. The recovered phase is then treated as if it were the true value of the carrier phase, and used in (partially) coherent detection of subsequent symbols. The work in [5], [6] follows this structured approach. This structured approach thus separates the carrier phase estimation problem from the data detection problem. It has the disadvantage that the entire structured receiver (phase estimator and data detector) is not optimized with respect to the ultimate criterion of interest in communication, namely, minimum SEP. Even the problem of optimizing a structured receiver with respect to the minimum SEP criterion does not seem to have been addressed at all. The nonstructured approach of Kobayashi [1] performs simultaneous maximum likelihood (ML) estimation of the carrier phase and the data sequence, i.e., it assumes that the carrier phase estimation problem and the data detection problem cannot be separated. However, it is not clear whether simultaneous estimation of carrier phase and data sequence will lead to a receiver that is optimum with respect to the minimum SEP criterion. In fact, our work shows that it does not. Kobayashi did not obtain any explicit receiver structure in [1]. Only an iterative approach for approximate computation of the ML estimates is developed. Another point is that ML estimation of the entire data sequence in [1] is more applicable to the coded case than to the present uncoded case. For an uncoded sequence, it is desired to minimize the expected number of symbol errors made in detecting the sequence, and this is achieved by the SBS receiver optimized with respect to the minimum SEP criterion. For the case of a coded sequence [4], it is necessary to minimize the probability of deciding on the wrong sequence, and this is achieved by the ML sequence estimator.

Macchi and Scharf [10] also followed the simultaneous estimation approach of Kobayashi. They developed a Viterbi algorithm for ML estimation of carrier phase and (uncoded) data sequence, but no attempt was made to determine the explicit receiver structure.

In [3], Falconer and Salz employed the unstructured approach and the minimum SEP criterion and considered SBS detection of a data sequence. They included an unknown timing phase in addition to the unknown carrier phase, but

Paper approved by E. Biglieri, the Editor for Data Communications and Modulation of the IEEE Communications Society. Manuscript received February 22, 1991.

The authors are with the Department of Electrical Engineering, National University of Singapore, Singapore 0511, Singapore.

IEEE Log Number 9401934.

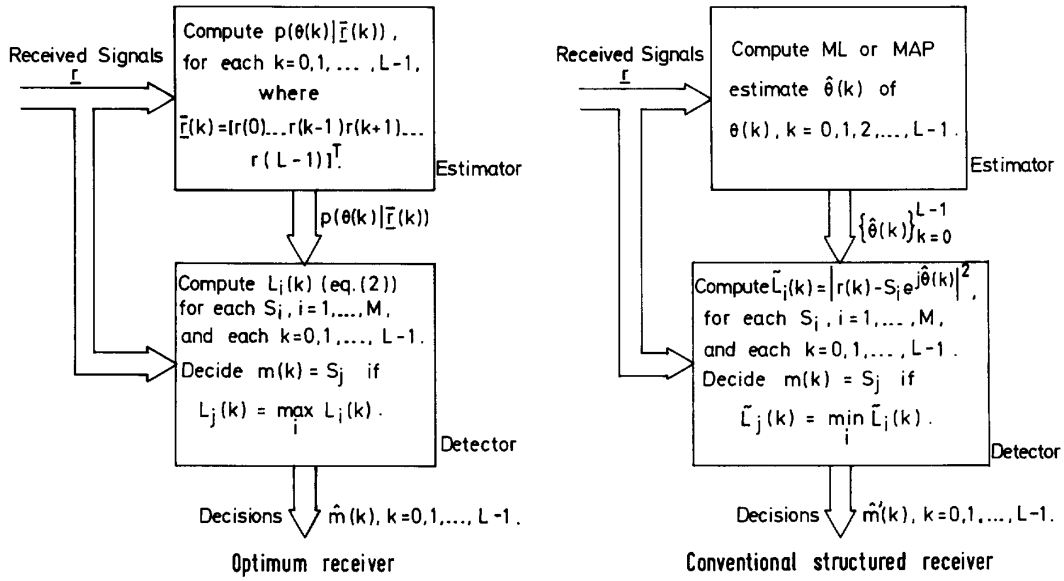


Fig. 1. Receiver structures.

explicit receiver results were obtained only for very high SNR's (signal-to-noise ratios). Because of the complexity of their results, it is not clear from [3] what systematic procedure a system designer can use to come up with a minimum SEP SBS receiver.

This paper obtains the explicit structure of the minimum SEP SBS receiver, and shows that it has a detector-estimator structure. In detecting a particular symbol, the totality of signals received over all the remaining symbol intervals are used in the estimator to compute the *a posteriori* or conditional pdf (probability density function) of the carrier phase in the interval concerned. The entire information concerning the carrier phase summarized in this conditional pdf is then employed in detecting the particular symbol. For lack of a better term, we also refer to the processor that computes the conditional pdf of carrier phase as an estimator. The optimum SBS receiver has a structure similar to the conventional structured receiver (see Fig. 1). The only difference is that the latter makes use of only an ML or MAP (maximum a posterior probability) estimate of the carrier phase as if it were the true value of the carrier phase in (partially) coherent data detection. This difference points out how the conventional structured receiver is suboptimum. The optimum receiver result also shows that the carrier phase estimation function and the symbol detection function are actually separable, i.e., the simultaneous estimation approach of Kobayashi [1] is wrong. Falconer and Salz [3] also conclude that the two functions are separable, but they have to assume very high SNR.

The optimum SBS receiver is, of course, nonimplementable in general, because computing the conditional pdf of the carrier phase is an infinite dimensional problem. The structure of the detector, and in fact the entire receiver, depend mainly on this conditional pdf of carrier phase. This latter pdf in turn depends primarily on the model of the unknown carrier phase. The key role played by the carrier phase model in the optimum SBS receiver design problem is thus highlighted. For most models of carrier phase in practice, it can be shown that this conditional pdf cannot be obtained in a simple closed form, a form which allows the explicit structure of the estimator and the detector to be specified. The only general way to deal with the problem of computing this conditional pdf is to represent the latter using a Fourier series as in [8] or some other suitable series. The estimator's function in computing the pdf is now to compute the Fourier coefficients. The explicit structure of the detector can also be specified in terms of these Fourier coefficients, as we will show. Thus, in principle, the Fourier series representation approach is one general way by which the optimum SBS receiver structure can be obtained explicitly. The implementation problem, however, is still infinite dimensional, because a very large (in theory, infinite) number of Fourier coefficients have to be computed each time in detecting a particular symbol. It is easy to see that as far as implementation is concerned, the problem is much simplified if the conditional pdf of carrier phase has a simple closed form, such as a Tikhonov pdf [9], which is completely characterized by a finite number of parameters. The computation of the conditional pdf by the estimator will, in this case, reduce to only the computation of these parameters, which will also completely characterize the detector. We will find that, indeed, there is one special case in which the conditional pdf of carrier phase is approximately Tikhonov, and in this case the estimator and detector parts of the receiver are simple and readily implementable. This is the case in which the phase process is slowly time-varying so that it can be considered constant at least over the duration of, say, $K + 1$ symbols, and we assume the SNR is sufficiently high. The optimum SBS receiver is then approximately a DA (decision-aided) receiver similar to those considered in [6], [11]. This result leads to a new interpretation for the latter receivers, namely, as optimum SBS receivers, and this interpretation applies also in particular to the concept of differential detection. Whether other models of carrier phase exist for which the conditional pdf has a simple closed form such as a Tikhonov pdf or some other canonical pdf's is an open question for further investigation.

The optimum SBS receiver, of course, reduces to the conventional structured receiver when the SNR is so high that the conditional pdf of carrier phase θ is very peaked around the MAP estimate $\hat{\theta}$ of θ so that this pdf is approximately an impulse. This provides justification for using the conventional structured receivers in practice which are easier to implement.

Having clarified the optimum SBS receiver structure, we next turn to its SEP performance. In general, the SEP of the receiver cannot be determined analytically. The only case in which the SEP can be obtained analytically is when the carrier phase can be considered constant over some $K+1$ symbol intervals, the SNR is sufficiently high so that we can employ decision-feedback, and the signal constellation is circular, such as the PSK constellations. This is one case, as stated previously, in which the explicit structure of the optimum SBS receiver can be obtained and readily implemented. Even in this case, the SEP can be analyzed only in the absence of decision errors, or in the presence of a known number of decision errors. We have done this in [6]. In the presence of a random number of decision errors, as is the case in practice, or in the presence of a rapidly varying phase, the SEP can only be obtained via computer simulations. We have simulated this DA receiver for 2, 4, and 8 PSK, not only for a constant carrier phase, but also for a Gaussian random-walk phase model and a model of a linearly increasing phase. The simulation results show good SEP performance for this DA receiver. The reader should note that this DA receiver, which assumes that the carrier phase is constant over some $K+1$ symbol intervals and ignores the details in the statistical model of the carrier phase process, is actually suboptimum for the Gaussian random-walk phase model and the linearly increasing phase model. The simulated SEP performance results, thus, indicate that even if the optimum SBS receivers for the latter phase models were implemented, the performance gains over this DA receiver would not be significant enough to justify the complexity involved in their implementations. This DA receiver is, therefore, an ideal approximation to the truly optimum SBS receiver for various phase models of practical interest.

The simulation studies showed that the DA receiver can encounter a "runaway", especially at low SNR. A variation of the DA receiver for PSK modulations is thus developed which overcomes this "runaway" problem. This receiver has the same structure as the DA receiver, and uses tentative decisions concerning future symbols fed forward, instead of firm past decisions fed back. The tentative decisions are made possible through the use of differential phase encoding. The decision feedforward receiver provides a stable technique for achieving coherent detection, and does not require any preamble for initial carrier phase acquisition, which makes it ideally suited to such applications as burst mode TDMA communications.

Section II presents the theory of minimum SEP SBS detection of an uncoded sequence, and Section III presents the simulated performance of the decision-feedback and decision-feedforward receivers for PSK signals.

II. Optimum Symbol-By-Symbol Detection

The received signal over the kth symbol interval is given by [6]

$$r(k) = m(k)e^{j\theta(k)} + n(k). \tag{1}$$

Here, $\boldsymbol{m} = [m(0)\ m(1)\ \cdots\ m(L-1)]^T$ is the sequence of uncoded data symbols transmitted, and $\boldsymbol{\theta} = [\theta(0)\ \theta(1)\ \cdots\ \theta(L-1)]^T$ is the unknown random carrier phase process. (Superscript T denotes transposed, and $*$ denotes complex conjugate). The latter is assumed to fluctuate slowly compared to the symbol rate so that the above piecewise-constant approximation can be made. The channel AWGN is $\boldsymbol{n} = [n(0)\ n(1)\ \cdots\ n(L-1)]^T$, a vector of independent and identically distributed (iid) complex Gaussian random variables (CGRV) with $E[\boldsymbol{n}] = \boldsymbol{0}$ and $E[\boldsymbol{n}\boldsymbol{n}^{*T}] = N_o\boldsymbol{I}$. We assume that $\boldsymbol{m}, \boldsymbol{n}$ and $\boldsymbol{\theta}$ are independent of one another. Because the data is uncoded, the symbols of $\boldsymbol{m}$ are independent, and each symbol can assume with equal probability any point $S_i, i = 1, \cdots, M$, in the signal constellation. The components of $\boldsymbol{\theta}$ are, however, correlated in general. The problem is to determine each symbol $m(k), k = 0, 1, \cdots, L-1$, with minimum SEP based on the totality of received signals $\boldsymbol{r} = [r(0)\ r(1)\cdots r(L-1)]^T$.

As is well-known [7], the optimum SBS receiver for the kth symbol $m(k)$ computes $L_i(k) = p(\boldsymbol{r}|m(k) = S_i)$ for each point S_i, $i = 1, \cdots, M$, in the signal constellation, and declares the decision $\hat{m}(k) = S_j$ if $L_j(k) = \max_i L_i(k)$. We will first show that

$$L_i(k) = C\int_{-\pi}^{\pi} p(r(k)|m(k) = S_i,\, \theta(k)) \cdot p(\theta(k)|\bar{\boldsymbol{r}}(k))\, d\theta(k) \tag{2}$$

where C is a constant independent of S_i, and $\bar{\boldsymbol{r}}(k) = [r(0)\ r(1)\ \cdots\ r(k-1)\ r(k+1)\ \cdots\ r(L-1)]^T$, i.e., $\bar{\boldsymbol{r}}(k)$ is the totality of signals received outside of the kth interval. The key steps in the derivation of (2) are as follows. We first write

$$L_i(k) = \int_{-\pi}^{\pi} p(\boldsymbol{r}|m(k) = S_i,\, \theta(k)) \cdot p(\theta(k)|m(k) = S_i)\, d\theta(k). \tag{3}$$

Next, we have

$$\begin{aligned} p(\boldsymbol{r}|m(k) = S_i,\, \theta(k)) &= p(r(k)|m(k) = S_i,\, \theta(k),\, \bar{\boldsymbol{r}}(k)) \\ &\quad \cdot p(\bar{\boldsymbol{r}}(k)|m(k) = S_i,\, \theta(k)) \\ &= p(r(k)|m(k) = S_i,\, \theta(k)) \\ &\quad \cdot p(\bar{\boldsymbol{r}}(k)|m(k) = S_i,\, \theta(k)), \end{aligned} \tag{4}$$

where we have used the fact that conditioned on $m(k)$ and $\boldsymbol{\theta}(k)$, the only randomness in $r(k)$ is due to $n(k)$, and this renders it independent of $\overline{\mathbf{r}}(k)$. Putting (4) into (3), we obtain (2) with $C = p(\overline{\mathbf{r}}(k))$ by noting that

$$\begin{aligned} &p(\overline{\mathbf{r}}(k)|m(k) = S_i,\, \theta(k))p(\theta(k)|m(k) = S_i) \\ &\quad = p(\overline{\mathbf{r}}(k),\, \theta(k)|m(k) = S_i) \\ &\quad = p(\theta(k)|\overline{\mathbf{r}}(k))p(\overline{\mathbf{r}}(k)), \end{aligned} \tag{5}$$

because $\theta(k)$ and $\overline{\mathbf{r}}(k)$ are both independent of $m(k)$.

From (1), we have

$$\begin{aligned} &p(r(k)|m(k) = S_i,\, \theta(k)) \\ &\qquad = \frac{1}{\pi N_0} \exp\left[\frac{-\left|r(k) - S_i e^{j\theta(k)}\right|^2}{N_0}\right]. \end{aligned} \tag{6}$$

The likelihood function $L_i(k)$ in (2) can now be evaluated in principle, and hence the structure of the optimal receiver can be specified explicitly, if the conditional pdf $p(\theta(k)|\overline{\mathbf{r}}(k))$ is known. Note that $\theta(k)$ depends on $\overline{\mathbf{r}}(k)$ because of the dependence among the components of $\boldsymbol{\theta}$. We will discuss later the computation of this pdf.

The detector-estimator structure of the optimum receiver is obvious from (2). For each time k, the estimator computes the conditional pdf of $\theta(k)$ given the signals $\overline{\mathbf{r}}(k)$ received outside the kth symbol interval. Suppose, for simplicity of discussion, that the phase $\theta(k)$ is a discrete random variable which takes on only one value in the set $\{\theta_l\}_{l=1}^{Q}$ for each time k. The estimator then computes the conditional probability $P_l(k) = P(\theta(k) = \theta_l|\overline{\mathbf{r}}(k))$ for each $l = 1, \cdots, Q$. The integral in (2) becomes a finite sum and the detector forms the weighted sum $L_i(k) = C(\pi N_o)^{-1} \sum_{l=1}^{Q} P_l(k) \exp[-|r(k) - S_i e^{j\theta_l}|^2/N_o]$, and makes the decision accordingly. This illustrates the roles played by the detector and the estimator in making a symbol decision. Note that the detector takes into account all possible values that $\theta(k)$ can take on, the weight given to each value θ_l being the conditional probability $P_l(k)$. The receiver structure is shown in Fig. 1.

The reader should note from the derivation of (2) how the data detection problem and the carrier phase estimation problem separate out. One key assumption is that the channel noise is AWGN, which makes the components of $\boldsymbol{n}$ independent. This renders the components of $\boldsymbol{r}$ independent when conditioned on given values of $\boldsymbol{m}$ and $\boldsymbol{\theta}$, and leads to the factor $p(r(k)|m(k) = S_i, \theta(k))$ in $L_i(k)$ [see (4)] which constitutes the detector part of the receiver. The other key assumptions are the independence between $\boldsymbol{m}$ and $\boldsymbol{\theta}$, and the independence among the uncoded symbols of $\boldsymbol{m}$. These latter assumptions lead to the term $p(\theta(k)|\overline{\mathbf{r}}(k))$ [see (5)] in $L_i(k)$ which constitutes the estimator part of the receiver. The independence between $\boldsymbol{m}$ and $\boldsymbol{\theta}$ in particular leads to $p(\theta(k)|\overline{\mathbf{r}}(k), m(k) = S_i) = p(\theta(k)|\overline{\mathbf{r}}(k))$ in (5), i.e., it separates the carrier phase estimation problem from the detection problem. This separation is important because it simplifies the receiver structure. The independence among the symbols of $\boldsymbol{m}$ leads to $p(\overline{\mathbf{r}}(k)|m(k) = S_i) = p(\overline{\mathbf{r}}(k))$, which gives the hypothesis-independent constant C in (2). If $\boldsymbol{m}$ were a coded sequence, this simplification would not result.

Consider now the pdf $p(\theta(k)|\overline{\mathbf{r}}(k))$ which will determine the structure of both the estimator part and the detector part of the receiver. In general, given the signal model (1) and a statistical model for the random process $\boldsymbol{\theta}$, it is impossible to analytically determine the pdf $p(\theta(k)|\overline{\mathbf{r}}(k))$. The most general way to deal with the problem is to represent the pdf using a Fourier series or some other suitable series. We will present here this general approach using the Fourier series representation. Since $\theta(k)$ is restricted to $[-\pi, \pi)$, $p(\theta(k)|\overline{\mathbf{r}}(k))$ is a periodic function when treated as a function of $\theta(k)$ over $(-\infty, \infty)$, and it admits the following Fourier series representation:

$$p(\theta(k)|\overline{\mathbf{r}}(k)) = \sum_{l=-\infty}^{\infty} c_l(k) e^{jl\theta(k)}, \quad -\pi \le \theta(k) < \pi. \tag{7}$$

Substituting (7) and (6) into (2) and integrating yields the result

$$\begin{aligned} L_i(k) &= CB \exp[-|S_i|^2/N_0] \\ &\quad \cdot \left\{ c_o(k) I_o(\beta_i(k)) + \sum_{l=1}^{\infty} (\pm 1)^l I_l(\beta_i(k)) \right. \\ &\qquad \left. \cdot [\cos(l\alpha_i(k)) a_l(k) + \sin(l\alpha_i(k)) b_l(k)] \right\}. \end{aligned} \tag{8}$$

Here, C is defined in (2), $B = (2/N_o) \exp\left[\frac{-|r(k)|^2}{N_o}\right]$ is independent of S_i, $\beta_i(k) = \frac{2}{N_o}|r(k)S_i^*|$, $\alpha_i(k) = \arg\left(\frac{2}{N_o} r(k) S_i^*\right)$, $a_l(k) = 2\,\mathrm{Re}\,[c_l(k)]$, $b_l(k) = -2\,\mathrm{Im}\,[c_l(k)]$, and $I_l(\cdot)$ is the modified Bessel function of the first kind of order l. The key step in deriving (8) is to write (6) as

$$\begin{aligned} &p(r(k)|m(k) = S_i,\, \theta(k)) \\ &\qquad = \frac{B}{2\pi} \exp\left[-\frac{|S_i|^2}{N_o} + \beta_i(k) \cos(\theta(k) - \alpha_i(k))\right], \end{aligned} \tag{9}$$

and then to expand the second exponential using the Jacobi–Anger formula [9, p. 100]:

$$e^{\pm\gamma \cos\phi} = I_o(\gamma) + 2\sum_{l=1}^{\infty} (\pm 1)^l I_l(\gamma) \cos l\phi, \tag{10}$$

before performing the integration in (2). The result (8) shows immediately the infinite-dimensional computational problem involved in implementing the detector even if the set of Fourier coefficients $\{c_l(k)\}_{l=-\infty}^{\infty}$ were available from the estimator. This, of course, severely limits the practical utility of the optimum receiver result in (8).

Since $p(\theta(k)|\overline{\mathbf{r}}(k))$ is represented in (7) as a Fourier series, the estimator's function now is to compute the coefficients $\{c_l(k)\}_{l=-\infty}^{\infty}$. This can be a very complicated task in general. We will present here one case in which these coefficients can be computed recursively in time k. In this case, assume for simplicity that $\boldsymbol{\theta}$ is given by the Gaussian random-walk model:

$$\theta(k+1) = \theta(k) + w(k) \tag{11}$$

where $\boldsymbol{w} = [w(0) \;\cdots\; w(L-2)]^T$ is a vector of iid Gaussian random variables with $E[\boldsymbol{w}] = \mathbf{0}$ and $E[\boldsymbol{w}\boldsymbol{w}^T] = \sigma^2 \boldsymbol{I}$. Also,

$\theta(0)$ is Gaussian with $E[\theta(0)] = 0$ and $E[\theta^2(0)] = \lambda^2$, and is independent of $\boldsymbol{w}$. The model (11) is a commonly employed model for random carrier phase [10]. We next make the receiver suboptimal by constraining it to detect the symbols sequentially in time k so that the decision on each symbol $m(k)$ is based only on the past and present received signals $[r(0)\ r(1)\ \cdots\ r(k-1)\ r(k)]^T$. Then, it is not difficult to see that the likelihood function $L_i(k)$ in (2) remains the same except that the conditional pdf $p(\theta(k)|\bar{\boldsymbol{r}}(k))$ is replaced by $p(\theta(k)|\boldsymbol{r}_p(k))$ where $\boldsymbol{r}_p(k) = [r(0)\ r(1)\ \cdots\ r(k-1)]^T$ is the set of past received signals at time k. We also assume the use of decision feedback, i.e., we assume all the past decisions $[\hat{m}(0)\ \hat{m}(1)\ \cdots\ \hat{m}(k-1)]^T$ are correct $(\hat{m}(l) = m(l))$ and that the past received signals are given by

$$r(l) = \hat{m}(l)e^{j\theta(l)} + n(l), \quad l = 0, 1, \cdots, k-1. \tag{12}$$

Using these assumptions and following the work of Willsky [8], it is now straightforward to show that $p(\theta(k)|\boldsymbol{r}_p(k))$ can be computed recursively in time k by recursively computing the Fourier coefficients $\{c_l(k)\}_{l=-\infty}^{\infty}$. The procedure is to first compute $p(\theta(k)|\boldsymbol{r}_p(k+1))$ starting from $p(\theta(k)|\boldsymbol{r}_p(k))$, and then to compute $p(\theta(k+1)|\boldsymbol{r}_p(k+1))$ starting from $p(\theta(k)|\boldsymbol{r}_p(k+1))$. The reader can refer to [8] for details. Note that although we assume the model (11), the recursive computation procedure here can be applied for a general Gauss–Markov model of $\boldsymbol{\theta}$.

It is obvious from [8] that the computational load imposed on the estimator is infinite dimensional even for this DA suboptimal receiver. This, coupled with the computational load of the detector (8), makes it extremely unlikely that any system designer would implement the optimal receiver in this most general form in practice.

For the optimal SBS receiver in (2) to be readily implementable, the conditional pdf $p(\theta(k)|\bar{\boldsymbol{r}}(k))$ has to have a suitable analytical form for which the integral in (2) can be evaluated. It will be ideal if this analytical form is characterized by a finite number of parameters so that computing the pdf in real time amounts to only computing these parameters. We have found one case in which $p(\theta(k)|\bar{\boldsymbol{r}}(k))$ has such an analytical form. In this case, we assume that θ is slowly time-varying so that it can be considered constant over a duration of at least $K+1$ symbol intervals where $K \ll L$. (Assume a very long sequence so that L is very large.) As in the Fourier series case, we constrain the receiver to detect the symbols sequentially in time k, and to be DA with perfect decision feedback so that the received signals are given by the model (12). In addition, we constrain the receiver to basing its decision concerning the kth symbol $m(k)$ on the present and the past K received signals $[r(k-K)\ \cdots\ r(k-1)\ r(k)]^T$. Then, again, the likelihood function $L_i(k)$ in (2) remains the same, except that the pdf $p(\theta(k)|\bar{\boldsymbol{r}}(k))$ is replaced by $p(\theta(k)|\boldsymbol{r}_p(k, K))$, where $\boldsymbol{r}_p(k, K) = [r(k-K)\ \cdots\ r(k-1)]^T$ is the set of received signals over the past K symbol intervals at time k. Letting θ denote the common value of $\theta(k-K), \cdots, \theta(k)$, and assuming θ to be uniformly distributed over $[-\pi, \pi)$, it is easy to show (see Appendix) that

$$p(\theta|\boldsymbol{r}_p(k, K)) = \frac{\exp\left[2N_o^{-1}|v(k)|\cos(\theta - \phi(k))\right]}{2\pi I_o(2N_o^{-1}|v(k)|)}. \tag{13}$$

This conditional pdf is a Tikhonov pdf centered at the mean value $\phi(k) = \arg\left(\frac{2}{N_o}v(k)\right)$ where

$$v(k) = \sum_{l=k-K}^{k-1} r(l)\hat{m}^*(l). \tag{14}$$

Using (13) and (6) in (2), and ignoring constants independent of S_i, we obtain

$$L_i(k) \sim \exp\left[-\frac{|S_i|^2}{N_o}\right] I_o\left(\frac{2}{N_o}|v(k) + r(k)S_i^*|\right). \tag{15}$$

The result (15) defines the suboptimal DA SBS receiver which bases its decision on the present and the past K received signals. For PSK modulations where $|S_i|^2 = E_s$, the energy per symbol for all i, $L_i(k)$ in (15) reduces to $L_i'(k)$ where

$$L_i'(k) \sim \mathrm{Re}\left[r(k)S_i^*v^*(k)\right]. \tag{16}$$

This receiver (16) is identical to the structured receiver in [6] in which a DA ML estimate of carrier phase is made from the past K received signals at each time k and then used in data detection as if it were the true value of carrier phase. In fact, we have shown in [6] that for PSK modulations, forming the coherent reference $v(k)$ in (14) is equivalent to making a DA ML estimate of the unknown carrier phase. The work here, thus, shows that the latter structured receiver of [6] is optimum if the receiver is constrained to be DA and to basing its decision concerning each symbol on the present and the past K received signals, if the carrier phase is indeed constant over $K+1$ symbol intervals. It also shows that the latter structured receiver makes optimum use of information concerning unknown carrier phase gleaned from the past K received signals. In [6], we have shown that the receiver (16) is a generalized differentially coherent receiver. By specializing (16) to the case of $K = 1$, we can also conclude that differential detection of DPSK makes optimum use of carrier phase information given by the previous received signal. As in [6], the decision-feedback assumption we made in arriving at (15) can be shown to be optimum at high SNR.

Note from (15) that, in general, explicit computation of the estimate of carrier phase is not required. The receiver achieves (partially) coherent detection through the establishment of the coherent reference $v(k)$ from the past K received signals. In the high SNR limit, since $I_o(x)$ behaves basically as e^x for $x \gg 1$, the result (15) reduces to

$$\begin{aligned}\Lambda_i(k) &= \ln L_i(k)\\ &\sim |v(k) + r(k)S_i^*| - \tfrac{1}{2}|S_i|^2.\end{aligned} \tag{17}$$

In this form, the receiver is readily implementable, since it requires only a square-root nonlinearity, which can be stored in a ROM, for instance. For PSK modulations, no nonlinearity is required. The SBS DA receiver (17) is identical to that arrived at in [11, eq. (8)] via sequence estimation.

Note that treating a slowly varying carrier phase as being constant over some $K+1$ symbol intervals is only a convenient approximation by which an implementable receiver can be obtained. The simulated performance results in Section III, however, will show that the approximation works well for phase models of practical interest.

III. Performance Results: Decision-Feedback and Decision-Feedforward Receivers

The error performance of the DA receiver (16) has been determined via computer simulations using various models of the phase process $\boldsymbol{\theta}$. We present the bit error probability (BEP) P_b results of the receiver (16) in Figs. 2–4, for 2, 4, and 8 PSK, respectively, as a function of SNR E_b/N_o where $E_b = E_s = |m(k)|^2$ for 2PSK, $E_b = E_s/2$ for 4 PSK, and $E_b = E_s/3$ for 8 PSK. Gray coding of bits onto signal points is assumed for 4 and 8 PSK. The Gaussian random-walk phase model refers to the model in (11), while the linearly increasing phase model is also given by (11) except that the inputs $\boldsymbol{w}$ are known constant phase increments, i.e., $w(k) = \Delta\theta, k = 0, \cdots, L-2$. Results for both actual decision-feedback (detected symbols fed back) and ideal decision feedback (no decision errors) are presented to show the performance loss due to decision errors. The simulation results show that the effect of decision errors on the receiver performance becomes negligible when the error probability is in the region of 10^{-3} or less. Thus, for the range of SNR E_b/N_o shown, the performance difference between actual decision feedback and ideal decision feedback is not noticeable for 2PSK, is slightly more pronounced for 4PSK, and is most significant for 8PSK. In general, it is also observed that the performance loss due to decision errors is higher for faster carrier phase variations, i.e., for a larger σ (the rms phase fluctuation in the random-walk model) or a larger $\Delta\theta$ (the constant phase increment in the linearly increasing model). The interval KT (T = symbol duration) over which the received signals are averaged via (14) to obtain the reference $v(k)$ is important as $(\mathrm{KT})^{-1}$ is a measure of the DA ML carrier phase estimator's bandwidth. The results of Fig. 5 for 2PSK with the Gaussian random-walk phase model show that a more rapidly varying phase requires an estimator of larger bandwidth $(\mathrm{KT})^{-1}$, or shorter averaging interval KT. Also, since the receiver is a DA receiver, we had to prevent a "runaway" due to decision errors by sending known symbols periodically so as to restart the estimator (14) with known symbols replacing the decisions $\hat{m}(l)$. The effect of this restart period on the receiver performance is demonstrated in Fig. 6 for 2PSK with the Gaussian random-walk phase model. Note also that for a constant unknown carrier phase, performance very close to that of coherent is obtained in all cases.

If the carrier phase is constant, the BEP P_b of the receiver (16) for 2PSK with no prior decision errors in the coherent reference $v(k)$ in (14) is given by [6, eq. (14)]:

$$P_b \approx \tfrac{1}{2}(1-K^{-1})^{-1/2}\,\mathrm{erfc}\,[E_b/N_o]^{1/2}, \quad K \geq 2, \quad (18a)$$

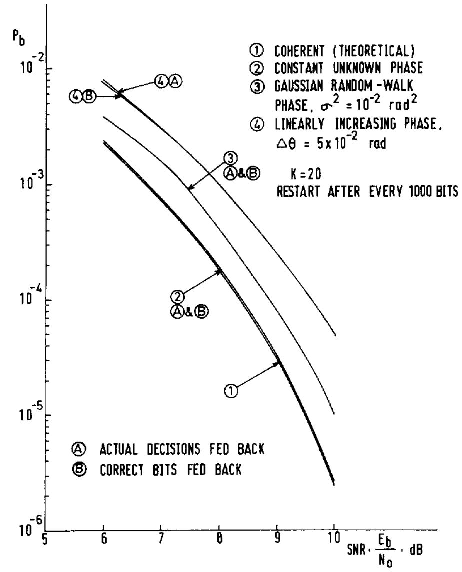


Fig. 2. Simulated BEP results of DA receiver for 2PSK.

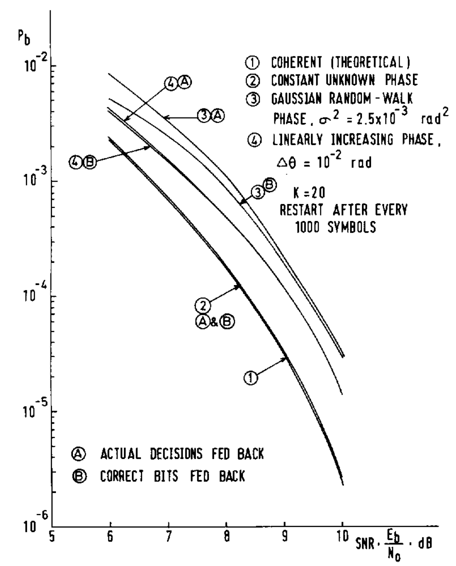


Fig. 3. Simulated BEP results of DA receiver for 4PSK.

for high SNR, i.e., $E_b/N_o \gg 1$. For $K = 1$, the receiver (16) is the ordinary differential detector, and its BEP P_b is

$$P_b = \tfrac{1}{2}\exp\,[-E_b/N_o]. \quad (18b)$$

By plotting P_b in (18) against E_b/N_o for different values of K the reader can easily show that most of the performance gain of (perfectly) coherent detection over differential detection can be recovered by merely the inclusion, in the receiver (16), of a memory of $K = 3$ or 4 past received signals for forming the reference $v(k)$. (See Fig. 7, for instance.) For larger values of K, the law of diminishing marginal returns sets in. A similar result can be shown to hold for MPSK in general. However, in practice, a larger value of K should be used so that the

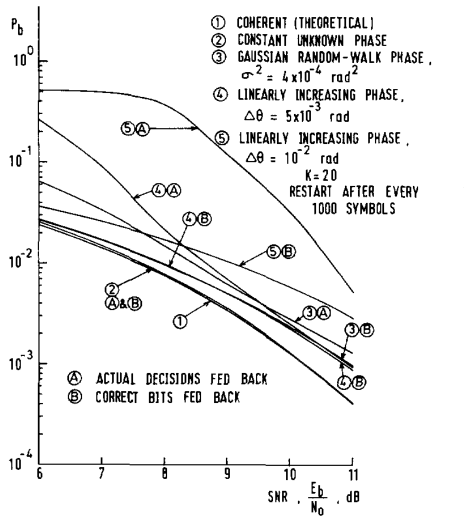


Fig. 4. Simulated BEP results of DA receiver for 8PSK.

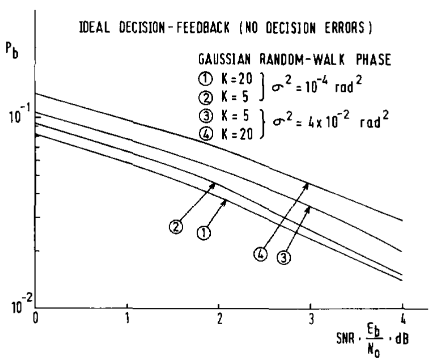


Fig. 5. Simulated BEP results of DA receiver for 2PSK with different averaging intervals.

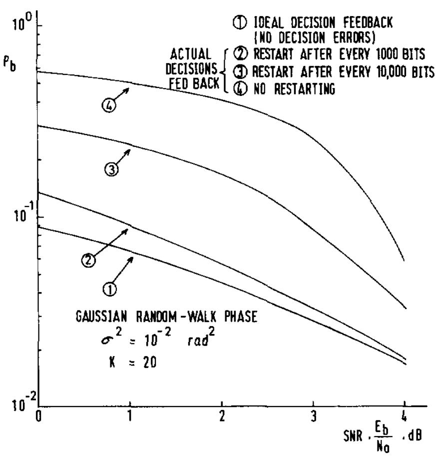


Fig. 6. Simulated BEP results of DA receiver for 2PSK with different restart intervals.

effect of prior decision errors on the receiver's performance can be minimized. This point has been proven by our analysis

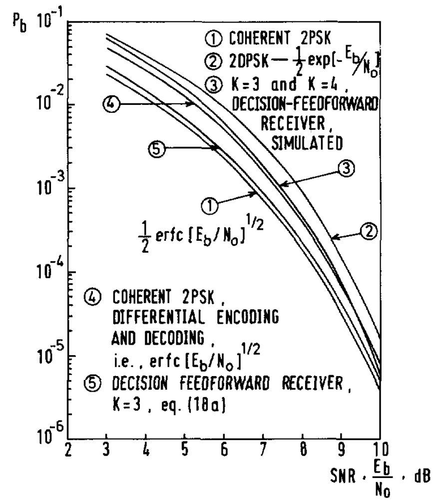


Fig. 7. Simulated BEP results of decision-feedforward receiver for 2DPSK.

in [6], and it explains our choice of the values of K in the simulation studies above.

The disadvantage of the DA receiver (16) is that it can suffer a "runaway" due to a burst of decision errors, especially at low SNR. To eliminate the possibility of a "runaway," we developed a modified receiver in which the coherent reference is formed through the use of feedforward of tentative decisions concerning future symbols. The carrier phase is still assumed constant over a duration of at least K symbol intervals. To enable tentative decisions to be made, we employ differential encoding of PSK (DPSK). Each symbol $m(k)$ is given by $m(k) = E_s^{1/2} e^{j\phi(k)}$. The information is encoded in the phase difference $\Delta\phi(k+1) = \phi(k+1) - \phi(k)$ ($m(0) = E_s^{1/2} e^{j\phi(0)}$ is the initial reference). To detect the information in $\Delta\phi(k+1)$, the receiver also employs generalized differentially coherent detection as in (16), i.e., it computes

$$q_i(k) = \mathrm{Re}\,[r(k) S_i^* \overline{v}^*(k)], \tag{19}$$

for each possible value $S_i, i = 1, \cdots, M$, of $e^{j\Delta\phi(k+1)}$, and decides that $\Delta\phi(k+1) = \arg(S_j)$ if $q_j(k) = \max_i q_i(k)$. The coherent reference $\overline{v}(k)$ in (19) is given by

$$\overline{v}(k) = z(k+1) \tag{20}$$

where the sequence of coherent references $[z(k+1)\ z(k+2)\ \cdots\ z(k+K)]^T$ is formed recursively as follows:

$$z(k+K) = r(k+K),$$

$$z(k+K-1) = r(k+K-1) + z(k+K) e^{-j\Delta\hat{\phi}(k+K)},$$

$$z(k+K-2) = r(k+K-2) + z(k+K-1) e^{-j\Delta\hat{\phi}(k+K-1)},$$

$$\vdots$$

$$z(k+1) = r(k+1) + z(k+2) e^{-j\Delta\hat{\phi}(k+2)}.$$

Here, $\Delta\hat{\phi}(k+l), l = 2, \cdots, K$, is the tentative decision on the phase information $\Delta\phi(k+l)$ obtained from the generalized differential detector similar to (19) which computes the decision statistic $\mathrm{Re}\,[r(k+l-1)S_i^* z^*(k+l)]$ for each possible value $S_i, i = 1, \cdots, M$, of $e^{j\Delta\phi(k+l)}$, and declares that $\Delta\hat{\phi}(k+l) = \arg(S_j)$ if S_j has the largest decision statistic. Thus, the reference $\overline{v}(k)$ is obtained through a process of progressively aligning the signal components of the received signals $r(k+K), r(k+K-1), \cdots$, and $r(k+1)$, until they are all in the same direction as the signal component $m(k+1)e^{j\theta(k+1)}$ of $r(k+1)$, assuming that all the tentative decisions $\{\Delta\hat{\phi}(k+l)\}_{l=2}^{K}$ are correct. The tentative decisions are formed in the reverse order in time, i.e., first $\Delta\hat{\phi}(k+K)$, then $\Delta\hat{\phi}(k+K-1)$, then $\Delta\hat{\phi}(k+K-2)$, and so on, similar to the order: first $z(k+K)$, then $z(k+K-1)$, then $z(k+K-2)$, and so on, for the coherent references. The tentative decisions are thus made with "stronger" and "stronger" coherent references, stronger in the sense of the SNR of the reference vector [6]. Thus, these decisions have lower and lower error probability, and the final decision on $\Delta\phi(k+1)$ has the smallest error probability.

If one uses only $K = 1$ future received signal in forming the reference $\overline{v}(k)$, the receiver (19) reduces to the ordinary differential detector. The decision feedforward receiver (19) is identical to the decision feedback receiver (16), except for the manner in which the coherent reference $\overline{v}(k)$ is formed. Note that differential encoding is employed in the receiver (19) only to enable tentative decisions to be made and fed forward to form the coherent reference $\overline{v}(k)$. Once $\overline{v}(k)$ is obtained, data detection in (19) is identical to that in (16), even though differential encoding is employed in the former. Thus, the two receivers (16) and (19) have the same error probability performance. In particular, for 2DPSK, the BEP P_b of the receiver (19) in the absence of tentative decision errors is given by [6, eq. (14)] or (18). As K tends to infinity, it is obvious from (18a) that the performance of the receiver (19) converges to that of coherent 2PSK. For $M \geq 4$, the same conclusion follows by letting K go to infinity in the symbol error probability result [6, eq. (13)]. Since the SNR of the coherent reference $\overline{v}(k)$ becomes infinite as K goes to infinity in the absence of tentative decision errors, it is easy to see why the performance of the receiver (19) converges to that of coherent MPSK. The effect of tentative decision errors on the receiver performance can only be assessed via computer simulations. The simulated BEP P_b results are presented in Figs. 7 and 8 for 2 and 4DPSK, respectively, with Gray coding assumed in the latter. In Fig. 7, we also plot the theoretical BEP (18a) for 2DPSK with $K = 3$ and with no tentative decision errors. The difference between this theoretical BEP curve and the simulation results shows the performance loss due to tentative decision errors. Note, as expected, that the performance loss decreases as the SNR increases, since for higher SNR the probability of tentative decision errors is smaller. Also, most of the performance gain of (perfectly) coherent detection over differential detection can be recovered by the receiver (19) with only a decision delay of $K = 3$ or 4 bit intervals for 2DPSK, and $K = 6$ or 7 symbol intervals for 4DPSK. This is similar to the previously stated

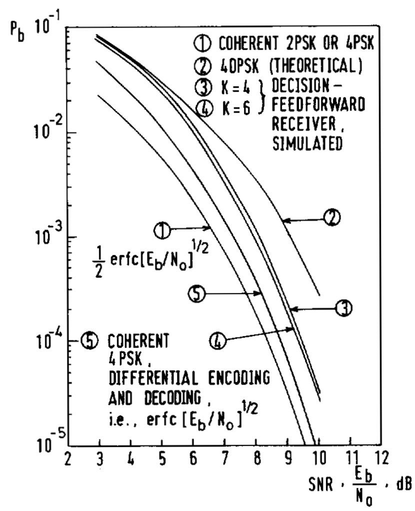


Fig. 8. Simulated BEP results of decision-feedforward receiver for 4DPSK.

result for the DA receiver (16), and is deduced theoretically from (18) for 2DPSK, and [6, eq. (13)] for 4DPSK. Our simulation studies have also confirmed this observation. For 2DPSK (Fig. 7), the simulated performance with $K = 4$ is not significantly improved from that with $K = 3$ except in the high SNR region. The same thing is observed for 4DPSK (Fig. 8) with values of K higher than 6. For both cases, the explanation is that the theoretical increment in performance gain due to a larger value of K is somewhat reduced by the larger number of tentative decision errors present within the K intervals, especially at lower SNR. For high SNR, the theoretical increment in performance gain due to a larger value of K is more closely realized in the simulations.

The effects of a time-varying carrier phase on the receiver performance are similar to those of the receiver (16), and will be omitted. The important advantage of the receiver (19) over the receiver (16) is that since the tentative decisions used for forming the coherent reference $\overline{v}(k)$ are made again each time a new reference is to be formed, there is no danger of the kind of error propagation which is encountered by the receiver (16). This has been confirmed in our simulation studies. Of course, the computational load on the receiver (19) is increased as compared to the receiver (16). However, this increase is only modest since in practice only a decision delay of $K = 3$ to 6 symbol intervals would be used, as mentioned above.

The receiver (19) is somewhat similar to the receiver based on multiple-symbol differential detection of MPSK developed by Divsalar and Simon [12]. The motivation of our receiver (19), however, differs from that in [12] which is based on sequence estimation. While [12] employs differential encoding to resolve ambiguity in sequence estimation, we employ differential encoding so that tentative decisions can be made and fed forward to enable a coherent reference to be formed. The receiver (19) can be regarded as a sequential, decision-feedforward, SBS version of the receiver of [12]. It has the advantage that it attains coherent MPSK performance as K

tends to infinity in the absence of tentative decision errors, i.e., to $\frac{1}{2}\operatorname{erfc}[E_b/N_o]^{1/2}$ for 2 and 4PSK, for instance. The performance of the receiver in [12], on the other hand, only converges to the performance of coherent MPSK with differential encoding and decoding, as the length K of the sequence being processed tends to infinity, i.e., to $\operatorname{erfc}[E_b/N_o]^{1/2}$ for 2 and 4PSK in the high SNR regime. Of course, the latter receiver is not affected by any prior decision errors, and the coherent performance just mentioned can be attained as K tends to infinity. The receiver (19), however, can also be expected to converge in performance to that of coherent MPSK (as K tends to infinity) in the high SNR regime where the tentative decision errors are rare. This behavior is, in fact, verified by the trends of the simulated performance curves of the receiver (19). As Fig. 7 shows for 2DPSK with $K = 3$, for instance, the simulated performance converges to the theoretical performance given by (18a) as the SNR increases. Thus, one can expect that for a sufficiently large K and sufficiently high SNR, the receiver (19) will perform better than coherent MPSK with differential encoding and decoding.

Edbauer [13] has also proposed a system employing differential MPSK and some K (> 1) differential detectors in the receiver, with the aim of achieving better than differential detection performance. His receiver is different from (19) in that at each time k it employs K differential detectors, each forming an output $r(k)r^*(k-l), l = 1, \cdots, K$, i.e., the receiver uses each of the past K received signals individually as a reference for $r(k)$. Using decision feedback to form a desired output for each detector, the receiver is optimized based on the heuristic approach of minimizing the quadratic errors of the outputs of these differential detectors. The basic receiver design approach as well as the resulting receiver structure are apparently different from ours.

IV. Conclusion

We have developed the theory for the design of the optimum SBS receiver for an uncoded data sequence received over the AWGN channel with unknown carrier phase. The optimum receiver structure is clarified, but in general it involves such a heavy computational load that it is very unlikely to be implemented in practice. One important special case has been identified in which the optimal receiver can be approximated by a readily implementable structure. The main contribution of this work lies in building a conceptual foundation for the optimum receiver design problem, a foundation which can be applied in other similar problems.

Appendix

To derive (13), first note that

$$p(\theta|\mathbf{r}_p(k, K)) = Gp(\mathbf{r}_p(k, K)|\theta) \tag{A1}$$

where $G = p(\theta)/p(\mathbf{r}_p(k, K))$ is independent of θ since $p(\theta) = 1/2\pi$, for $-\pi \le \theta < \pi$. In general, θ is *a priori* uniform over $[-\pi, \pi)$. Now, from the DA signal model (12), we have

$$p(\mathbf{r}_p(k, K)|\theta) = \prod_{l=k-K}^{k-1} p_{n(l)}(r(l) - \hat{m}(l)e^{j\theta}) = D\exp\left(\frac{2}{N_o}\sum_{l=k-K}^{k-1} \operatorname{Re}[r(l)\hat{m}^*(l)e^{-j\theta}]\right) \tag{A2}$$

where $D = (\pi N_o)^{-K}\exp[-(1/N_o)\sum_{l=k-K}^{k-1}(|r(l)|^2 + |m(l)|^2)]$. The exponent in (A2) can be written as $2N_o^{-1}\operatorname{Re}[v(k)e^{-j\theta}] = 2N_o^{-1}|v(k)|\cos(\theta - \phi(k))$. Putting (A2) into (A1) now gives (13). The constant GD can be shown to be $1/\{2\pi I_o(2N_o^{-1}|v(k)|)\}$ for normalization.

References

[1] H. Kobayashi, "Simultaneous adaptive estimation and decision algorithm for carrier modulated data transmission systems," *IEEE Trans. Commun.*, vol. COM-19, pp. 268–280, June 1971.

[2] W. C. Lindsey and M. K. Simon, *Telecommunication Systems Engineering.* Englewood Cliffs, NJ: Prentice-Hall, 1973.

[3] D. D. Falconer and J. Salz, "Optimal reception of digital data over the Gaussian channel with unknown delay and phase jitter," *IEEE Trans. Inform. Theory,* vol. IT-23, pp. 117–126, Jan. 1977.

[4] P. Y. Kam and P. Sinha, "Optimum detection of digital data over the Gaussian channel with unknown carrier phase—The coded case," submitted to *IEEE Trans. Commun.,*.

[5] A. J. Viterbi and A. M. Viterbi, "Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission," *IEEE Trans. Inform. Theory,* vol. IT-29, pp. 543–551, July 1983.

[6] P. Y. Kam, "Maximum likelihood carrier phase recovery for linear suppressed-carrier digital data modulations," *IEEE Trans. Commun.*, vol. COM-34, pp. 522–527, June 1986.

[7] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I.* New York: Wiley, 1968.

[8] A. S. Willsky, "Fourier series and estimation on the circle with applications to synchronous communication—Part I: Analysis," *IEEE Trans. Inform. Theory,* vol. IT-20, pp. 577–583, Sept. 1974.

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

[10] O. Macchi and L. L. Scharf, "A dynamic programming algorithm for phase estimation and data decoding on random phase channels," *IEEE Trans. Inform. Theory,* vol. IT-27, pp. 581–595, Sept. 1981.

[11] P. Y. Kam, "Maximum-likelihood digital data sequence estimation over the Gaussian channel with unknown carrier phase," *IEEE Trans. Commun.,* vol. COM-35, pp. 764–767, July 1987.

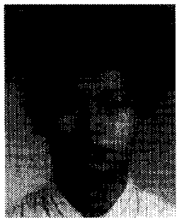
[12] D. Divsalar and M. K. Simon, "Multiple-symbol differential detection of MPSK," *IEEE Trans. Commun.,* vol. 38, p. 300–308, Mar. 1990.

[13] F. Edbauer, "Bit error rate of binary and quaternary DPSK signals with multiple differential feedback detection," *IEEE Trans. Commun.,* vol. 40, pp. 457–460, Mar. 1992.

Pooi Yuen Kam (M'83–SM'87) was born in Ipoh, Malaysia, on April 12, 1951. He received the S.B., S.M., and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1972, 1973, and 1976, respectively.

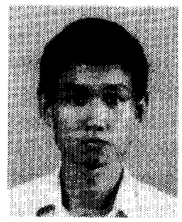
From 1976 to 1978 he was a member of the Technical Staff at Bell Laboratories, Holmdel, NJ, where he was involved in packet-network studies. Since 1978 he has been with the Department of Electrical Engineering, National University of Singapore, Kent Ridge, Republic of Singapore, where he has been teaching courses in communication theory and systems. His research interests are in stochastic processes, detection and estimation theory, and digital communications. During the academic year 1987–1988 he was a Hitachi Foundation Visiting Research Fellow at the Department of Physical Electronics, Tokyo Institute of Technology, Tokyo, Japan.

Dr. Kam is a member of Tau Beta Pi, Eta Kappa Nu, and Sigma Xi.



Seng Siew Ng was born in April 1964. He received the B.Eng. (Second Class Lower Honors) degree in electrical engineering from the National University of Singapore, Kent Ridge, in 1989.

From 1989 to 1990, he was a reliability engineer with AT&T Consumer Products Pte Ltd, Singapore. Since 1990, he has been a systems engineer with IBM Singapore Pte Ltd. Currently, he is working in the areas of systems network architecture and local area networks, in particular, the IBM token-ring network.



Tock Soon Ng was born in Singapore in May 1965. He received the B.Eng (First Class Honors) degree in electrical engineering from the National University of Singapore, Kent Ridge, in 1990.

Since 1990, he has been a systems engineer with IBM Singapore (Pte) Ltd, Public Sector Branch, working in the areas of IBM OS/2, mobile computing, and client/server solutions. In 1991, he was awarded the Manager Recognition Award by IBM Singapore for the implementation of a local area network prototype in the Overseas Mission Project.