

Improved Receivers for Coherent FSK Systems

Hongsheng Gao, Peter J. Smith, *Member, IEEE*, and Mansoor Shafi, *Fellow, IEEE*

Abstract—There has been considerable interest over the last decade in the detection of digital coherent lightwave signals corrupted by phase noise. However, the majority of the work has been in the performance analysis of various modulation schemes. In addition, the receivers studied have, in general, been zero-phase noise optimal or *ad hoc* modifications of these. In this paper, we consider the design problem of constructing optimal receivers for systems using frequency shift keying (FSK). Using the innovations approach, we show the format of the optimal receiver and use a small phase noise approximation to derive suboptimal receivers. Also, receivers based on optimum linear filters are also derived. Both receivers are analyzed by simulation and performance improvements over the standard receivers are shown. Furthermore, the receivers developed can equally well be used to further improve the performance of other detection schemes such as postdetection filtering and time diversity. The methodology is also used to derive equivalent receivers for systems using on-off keying (OOK).

Index Terms—Frequency shift keying, optical communication, receivers.

I. INTRODUCTION

THE performance of coherent optical systems in the presence of both additive shot noise and laser phase noise has been studied for around a decade. Some of the development of this work can be found in [1]–[17]. The current state of this research is that accurate performance evaluations can be made for a variety of modulation schemes using approaches which adhere fully to the Brownian motion model of phase noise. The pioneering work by Foschini *et al.* [1], [2] and Garret *et al.* [3]–[10] forms perhaps the cornerstone of all these developments. It should be pointed out, however, that analysis remains either highly computer intensive, technically challenging, or both! The nature of the phase noise variable after filtering during detection is too complex to permit a simple analysis.

In this paper, we consider a less well-studied problem which is that of receiver design. Most of the receivers studied in the literature are standard designs which are optimal in the absence of phase noise. Relatively little work has been done on designing receivers to combat the effects of phase noise. Foschini *et al.* [2] considered both shortening the bit period and using postdetection filtering for frequency shift keying (FSK) and on-off keying (OOK) systems. Their results showed that performance gains can be achieved by both approaches. Postdetection filtering has also been considered

by Jacobsen and Garret [6], [7]. On a different tack, Dallal and Shamai [11] used a time diversity approach to improving differential phase shift keying (DPSK) performance by the repeated transmission of a signal over several chips in one bit period. More recently, Kaiser *et al.* [12] used an *ad hoc* “weighting” filter to improve DPSK performance. The use of the innovations approach, developed by Kailath [18], [19], in heterodyne optical communication appears to have been pioneered by the excellent work of Georghiadis [22], [23], which appeared a decade ago. In this paper we use the same innovations approach in the study of FSK and OOK systems.

The layout of the paper is as follows. In Section II, we discuss the type of signals used in FSK and OOK and the noise variables present. We also show how Kailath’s innovations approach can be applied to construct a theoretical receiver form which is optimal in the maximum likelihood sense. The true maximum likelihood receiver depends on the derivation of a certain least squares estimator (LSE). Unfortunately the exact LSE is very difficult to find. Hence, in Section III, two approximate LSE’s are derived which are used in the optimal receiver form to give the improved receivers discussed in Section IV where the performance of these receivers is studied by simulation. Performance gains over standard receivers and the utility of these receivers is demonstrated.

II. SIGNAL, NOISE, AND RECEIVER MODEL

A. The Signal

Following the notation of Foschini *et al.* [2] the signals for every T -second data interval can be written in complex envelope notation as

FSK Modulation

$$r(t) = \begin{cases} S \exp(j(2\pi\Delta f t + \theta(t) + \phi)) + n(t) & \text{Data} = 1 \\ S \exp(j(-2\pi\Delta f t + \theta(t) + \phi)) + n(t) & \text{Data} = 0 \end{cases} \quad (1)$$

OOK Modulation

$$r(t) = \begin{cases} S \exp(j(\theta(t) + \phi)) + n(t) & \text{Data} = 1 \\ n(t) & \text{Data} = 0 \end{cases} \quad (2)$$

where S^2 is proportional to the power in the received optical pulse, $2\Delta f$ is the frequency separation, $\theta(t)$ is the composite phase noise due to the transmitting and receiving lasers, and $n(t)$ is a complex waveform representing additive noise at the photodetector output.

Manuscript received April 21, 1997; revised June 29, 1998.

H. Gao and P. J. Smith are with the Institute of Statistics and Operations Research, Victoria University of Wellington, Wellington, New Zealand.

M. Shafi is with Fundamental Planning, Telecom New Zealand Ltd., Wellington, New Zealand.

Publisher Item Identifier S 0733-8724(98)08299-1.

B. The Noise Processes

The phase noise process, $\theta'(t)$, is modeled as a Brownian motion process with $\theta'(0) = \phi$ where ϕ is a uniform random variable on $(-\pi, \pi]$. Throughout the paper we denote an arbitrary bit period by $[0, T]$. Hence, ϕ is the phase value at the start of a bit period. Defining $\theta(t) = \theta'(t) - \phi$ we can model $\theta(t)$ as a zero-mean Brownian motion process with $\text{Var}(\theta(t)) = 2\pi\beta t$. The additive shot noise, $n(t)$, is modeled as a complex Gaussian white noise process with two-sided spectral density N_0 . Two parameters sufficient to characterize the noise processes are $\zeta = \beta T$ and $\xi = S^2 T / N_0$.

C. Receiver Structure

Let $r(t)$ be a vector¹ random process on the interval $[0, T]$ of the form

$$r(t) = z(t) + n(t) \tag{3}$$

where $n(t)$ is WGN with unit spectral intensity

$$E[n(t)] = 0, \quad E[n(t)n^T(s)] = I\delta(t - s) \tag{4}$$

the superscript T denotes transpose, I is the identity matrix, and $z(t)$, which is called the signal process and is not necessarily a Gaussian process, obeys the following:

- the process $z(t)$ has finite expected energy over the interval $[0, T]$, such that

$$\int_0^T E[z(t)^T z(t)] dt < \infty \tag{5}$$

$$\int_0^T E[(z(t) - E[z(t)])^T (z(t) - E[z(t)])] dt < \infty. \tag{6}$$

- $r(t)$ may depend upon past $\{n(\tau), 0 \leq \tau < t\}$, but future WGN $n(t)$ is independent of past $r(\tau)$, i.e.,

$$E[n(t)r(\tau)^T] = 0, \quad 0 \leq \tau < t \leq T. \tag{7}$$

If we define $\hat{z}(t) \triangleq$ the causal least-squares estimate of $z(t)$ given $\{r(\tau), 0 \leq \tau < t < T\}$, then it is well known that

$$\hat{z}(t) = E[z(t)|r(\tau), 0 \leq \tau < t < T] \tag{8}$$

i.e., the conditional mean of $z(t)$ given that it is only possible to observe the noisy process $r(\tau)$ up to time t .

Consider the signal detection problem with hypotheses of the form

$$\begin{aligned} H_1: r(t) &= z(t) + n(t) & t \in [0, T] \\ H_0: r(t) &= n(t) & t \in [0, T] \end{aligned} \tag{9}$$

where $n(t)$ is a Gaussian noise process which satisfies (5) and (6) with $E[n(t)n^T(\tau)] = I\delta(t - \tau)$. The usual performance criteria is to minimize the bit error rate (BER) of the system, which can be achieved through a maximum likelihood decision between the two hypotheses. Kailath [18] showed the likelihood ratio (LR) can be written

$$\text{LR} = \exp \left\{ \int_0^T \hat{z}_1^T(t)r(t) dt - \frac{1}{2} \int_0^T \hat{z}_1^T(t)\hat{z}_1(t) dt \right\} \tag{10}$$

¹Note that we use vector notation instead of complex numbers since the innovations theory used later is commonly expressed in vector form.

where $\hat{z}_1(t) = E[z(t)|\{r(\tau), 0 \leq \tau < t\}, H_1]$, and \int denotes a special kind of integral called an $It\hat{o}$ stochastic integral. Since $\hat{z}_1(t)$ will itself be an integral, see (13), there is a stochastic double integral in (10) called a double Wiener- $It\hat{o}$ integral. The main feature of this integral is that its mean value is zero when $z(t) = 0$. A description of the integral and its computation is given in [20].

The fundamental problem is the derivation of $\hat{z}_1(t)$ since no general technique exists to find $\hat{z}_1(t)$. However, if the signal process $z(t)$ is Gaussian there is a solution. Define

$$K(t, s) \triangleq E[(z(t) - E[z(t)])(z(s) - E[z(s)])^T] \tag{11}$$

$$0 \leq s, t < T$$

and assume that $K(t, s)$ is continuous in (t, s) and

$$\int_0^T \text{tr}[K(t, t)] dt < \infty \tag{12}$$

then $\hat{z}(t)$ will be linear and

$$\hat{z}(t) = \int_0^t h(t, \tau)r(\tau) d\tau \tag{13}$$

where $h(t, s)$ must satisfy the well known Wiener-Hopf equation

$$h(t, s) + \int_0^t h(t, \tau)K(\tau, s) d\tau = K(t, s) \quad 0 \leq s < t < T. \tag{14}$$

The above ‘‘on-off’’ situation studied by Kailath is easily extended to the general binary detection problem where the receiver has to decide between two types of signal. Here the general hypotheses are

$$\begin{aligned} H_1: r(t) &= z_1(t) + n(t) & t \in [0, T] \\ H_2: r(t) &= z_2(t) + \nu(t) & t \in [0, T] \end{aligned} \tag{15}$$

where $n(t)$ and $\nu(t)$ are Gaussian noise processes and the signal processes $z_1(t)$ and $z_2(t)$ satisfy the finite power constraints (5) and (6). Note that all terms $r(t)$, $z_i(t)$, $n(t)$, and $\nu(t)$ can be scalar or vector quantities and this includes complex quantities as they can be expressed as a real and imaginary vector pair. It is trivial to show that the FSK and OOK signals (1) and (2) satisfy both the form of (15) and the constraints in (5) and (6). Hence, the maximum likelihood receiver structure derived by Kailath is applicable to both FSK and OOK systems. Using Kailath [19] and correcting a few small errors the maximum likelihood receiver for the detection problem (15) is given by

$$\begin{aligned} \text{MLR} &= \int_0^T (\hat{z}_1^T(t) - \hat{z}_2^T(t))r(t) dt \\ &\quad - \frac{1}{2} \int_0^T (\hat{z}_1^T(t)\hat{z}_1(t) - \hat{z}_2^T(t)\hat{z}_2(t)) dt \end{aligned} \tag{16}$$

where

$$\begin{aligned} \text{MLR} > 0 &\implies H_1 \\ \text{MLR} < 0 &\implies H_2 \end{aligned} \tag{17}$$

and $\hat{z}_1(t)$ and $\hat{z}_2(t)$ are the least squares estimators (LSE's) of $z_1(t)$ and $z_2(t)$ given by

$$\hat{z}_i(t) = E[z_i(t)|\{r(\tau), 0 \leq \tau < t\}, H_i], \quad i = 1, 2. \quad (18)$$

The power of the innovations approach is that the simple receiver format in (16) is completely general to all problems of the form in (15). As discussed in the above section, the price to be paid for such generality is the difficulty in computing the LSE's in (18), and only for Gaussian signal processes is there a known formula (13) for these. In this work therefore we begin by using a Gaussian approximation to the signals, compute the resulting LSE's and substitute these approximate LSE's in the exact receiver structure. In a later section we also compute the optimal linear LSE and use this in the exact receiver structure.

Since $r(t)$, $z_1(t)$, and $z_2(t)$ are complex quantities in our applications it is convenient to rewrite the vector form of (16) as below

$$\begin{aligned} \text{MLR} &= \int_0^T \text{Re}\{(\hat{z}_1(t) - \hat{z}_2(t))r^*(t)\} dt \\ &\quad - \frac{1}{2} \int_0^T (|\hat{z}_1(t)|^2 - |\hat{z}_2(t)|^2) dt \end{aligned} \quad (19)$$

where * denotes the complex conjugate of $r(t)$. Note that in (16) we express the terms as vectors and in (19) as complex quantities. Hence we only need to derive the approximate LSE's and the resulting receiver structure is given by (19).

III. IMPROVED RECEIVERS USING A SMALL PHASE LSE

A. Small Phase Approximation for the LSE

For both FSK and OOK and Data = 0 or 1 the signal can be written

$$r(t) = z_i(t) + n(t), \quad i = 1, 2 \quad (20)$$

where

$$z_i(t) = S_i \exp(j(\alpha_i t + \theta(t) + \phi)) \quad (21)$$

and $S_i \in \{0, S\}$, $\alpha_i = \pm 2\pi\Delta f$ or zero. Dropping the subscripts and the dependence on H_i for convenience we consider the LSE $\hat{z}(t)$ of the general $z(t)$ process in (21)

$$\begin{aligned} \hat{z}(t) &= \exp(j\alpha t) E_\phi[\exp(j\phi) E_\theta(S \exp(j\theta(t)) | \phi \\ &\quad r(s), 0 \leq s < t)] \\ &= \exp(j\alpha t) E_\phi[\exp(j\phi) E_\theta(S \exp(j\theta(t)) \\ &\quad - m(t) + m(t) | \phi, r(s), 0 \leq s < t)] \\ &= \exp(j\alpha t) E_\phi[\exp(j\phi) E_\theta(S \exp(j\theta(t)) \\ &\quad - m(t) | \phi, r(s), 0 \leq s < t)] \end{aligned}$$

since $m(t) = E[S \exp(j\theta(t))]$ is a known function and $E_\phi[\exp(j\phi)] = 0$. Now, we define $Z(t) \triangleq S \exp(j\theta(t)) - m(t)$ and the above equation becomes

$$\hat{z}(t) = \exp(j\alpha t) E_\phi[\exp(j\phi) E_\theta(Z(t) | \phi, R(s), 0 \leq s < t)] \quad (22)$$

where

$$R(s) = \exp(-j(\alpha s + \phi)) r(s) - m(s). \quad (23)$$

This is valid since α and $m(t)$ are known, we have conditioned on ϕ and hence the information in $r(s)$ and $R(s)$ is equivalent. Furthermore, $R(s)$ can be written

$$R(s) = Z(s) + \exp(-j(\alpha s + \phi)) n(s) \triangleq Z(s) + \bar{n}(s) \quad (24)$$

where it is trivial to show that $\bar{n}(s)$ is another complex noise process with the same statistics as $n(s)$.

Now, we are able to apply standard results to approximate the LSE in (22). Note that the inner expectation in (22) defines the LSE of $Z(t)$ given the observations $R(s) = Z(s) + \bar{n}(s)$ over $[0, t)$ and assuming ϕ is known. Denote this LSE as

$$\hat{Z}(t) = E_\theta[Z(t) | \phi, R(s) = Z(s) + \bar{n}(s), 0 \leq s < t]$$

where the complex process $R(s) = R_1(s) + jR_2(s)$ is written in vector form as

$$R(s) = \begin{bmatrix} R_1(s) \\ R_2(s) \end{bmatrix}.$$

Clearly $Z(t)$ is zero mean, but it is not Gaussian. Hence, at this stage, we introduce the small phase noise approximation

$$\exp(j\theta(t)) \approx 1 + j\theta(t). \quad (25)$$

Such small phase approximations have often been used and empirical evidence exists that they lead to accurate performance analysis for coherent optical systems with $\beta T \leq 0.1$ [2], [8], [13], [24]. How closely receivers based on such approximations match optimal receivers is essentially unknown since optimal receivers have not been found. Nevertheless, these approximations lead to a model for which optimal receivers can be developed and thus provide a sensible design strategy. With this approximation $m(t) = S$ and $Z(t) = jS\theta(t)$ is zero mean Gaussian. Hence, (13) can be used, i.e.,

$$\hat{Z}(t) = \int_0^t h(t, \tau) R(\tau) d\tau \quad (26)$$

where $h(t, s) = \begin{bmatrix} h_1(t, s) & h_2(t, s) \\ h_3(t, s) & h_4(t, s) \end{bmatrix}$ must satisfy the well-known Wiener-Hopf equation

$$h(t, s) + \int_0^t h(t, \tau) K(\tau, s) d\tau = K(t, s) \quad 0 \leq s < t < T.$$

$K(t, s)$ is the covariance matrix of $Z(t)$ given by

$$\begin{aligned} K(t, s) &= E\left\{ \begin{bmatrix} 0 \\ S\theta(t) \end{bmatrix} \begin{bmatrix} 0, S\theta(s) \end{bmatrix} \right\} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & S^2 \gamma \min(s, t) \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & S^2 \gamma s \end{bmatrix} \end{aligned} \quad (27)$$

where the notation $\gamma = 2\pi\beta$ is used. Hence, the above Wiener–Hopf equation becomes

$$\begin{aligned} & \begin{bmatrix} h_1(t, s) & h_2(t, s) \\ h_3(t, s) & h_4(t, s) \end{bmatrix} + \int_0^s \begin{bmatrix} h_1(t, \tau) & h_2(t, \tau) \\ h_3(t, \tau) & h_4(t, \tau) \end{bmatrix} \\ & \cdot \begin{bmatrix} 0 & 0 \\ 0 & S^2\gamma\tau \end{bmatrix} d\tau + \int_s^t \begin{bmatrix} h_1(t, \tau) & h_2(t, \tau) \\ h_3(t, \tau) & h_4(t, \tau) \end{bmatrix} \\ & \cdot \begin{bmatrix} 0 & 0 \\ 0 & S^2\gamma s \end{bmatrix} d\tau = \begin{bmatrix} 0 & 0 \\ 0 & S^2\gamma s \end{bmatrix}. \end{aligned} \quad (28)$$

The solution to (28) is $h_1(t, s) = h_2(t, s) = h_3(t, s) = 0$ and $h_4(t, s)$ is given by

$$\begin{aligned} h_4(t, s) + S^2\gamma \int_0^s \tau h_4(t, \tau) d\tau \\ + S^2\gamma s \int_s^t h_4(t, \tau) d\tau = S^2\gamma s. \end{aligned} \quad (29)$$

If we assume a separable solution to (29) of the form $h_4(t, s) = f(t)g(s)$ with series expansions $f(t) = \sum_{i=0}^{\infty} a_i t^i$, $g(s) = \sum_{i=0}^{\infty} b_i s^i$ then the simple approach of equating coefficients in (29) yields the solution

$$h_4(t, s) = w \left[\frac{e^{ws} - e^{-ws}}{e^{wt} + e^{-wt}} \right], \quad w = \gamma^{1/2} S. \quad (30)$$

Hence, from (26) the estimator $\hat{Z}(t)$ is given in complex notation by

$$\hat{Z}(t) = j \int_0^t h_4(t, s) R_2(s) ds. \quad (31)$$

From (22) the LSE $\hat{z}(t)$ can be written

$$\hat{z}(t) = \exp(j\alpha t) E_\phi \left[\exp(j\phi) j \int_0^t h_4(t, s) R_2(s) ds \right]. \quad (32)$$

If we set $r(t) = r_1(t) + jr_2(t)$, from (23) $R_2(s)$ is given by

$$R_2(s) = \cos(\alpha s + \phi) r_2(s) + \sin(\alpha s + \phi) r_1(s). \quad (33)$$

Substituting (33) in (32) and using the expectation results

$$\begin{aligned} E_\phi(\cos \phi) &= E_\phi(\sin \phi) = E_\phi(\sin \phi \cos \phi) = 0 \\ E_\phi(\cos^2 \phi) &= E_\phi(\sin^2 \phi) = \frac{1}{2} \end{aligned} \quad (34)$$

gives, after a little algebra

$$\hat{z}(t) = \frac{1}{2} \exp(j\alpha t) \int_0^t h_4(t, s) \exp(-j\alpha s) r(s) ds. \quad (35)$$

B. Improved Receivers

For FSK systems we have

$$\begin{aligned} H_1: \text{Data} = 1 \\ H_2: \text{Data} = 0 \end{aligned} \quad (36)$$

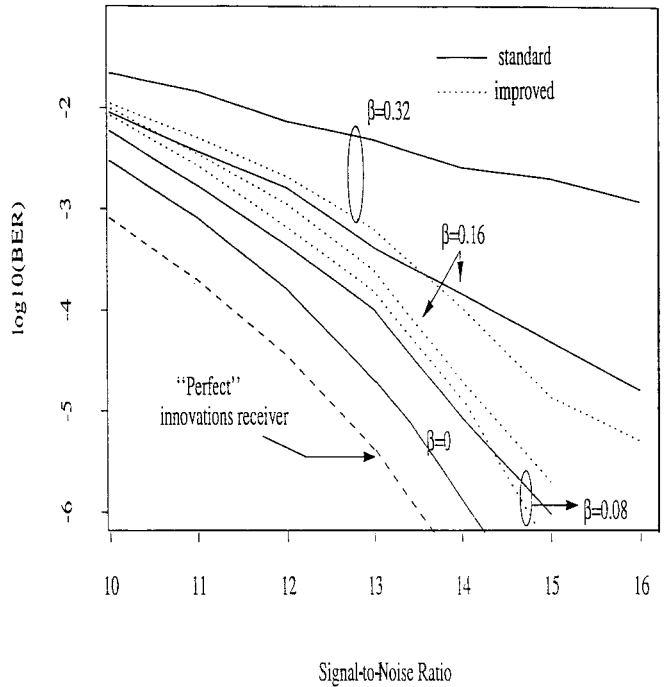


Fig. 1. Improved receiver using a small phase LSE.

or

$$\begin{aligned} H_1: \alpha_1 = 2\pi\Delta f, \quad S_1 = S \\ H_2: \alpha_2 = -2\pi\Delta f, \quad S_2 = S. \end{aligned} \quad (37)$$

Substituting in (35) gives

$$\begin{aligned} \hat{z}_1(t) &= \frac{1}{2} \exp(j2\pi\Delta ft) \int_0^t h_4(t, s) \exp(-j2\pi\Delta fs) r(s) ds \\ \hat{z}_2(t) &= \frac{1}{2} \exp(-j2\pi\Delta ft) \int_0^t h_4(t, s) \exp(j2\pi\Delta fs) r(s) ds. \end{aligned} \quad (38)$$

Hence, from (19), the improved receiver has the form below, after a little algebra

$$\begin{aligned} \text{MLR} &= \frac{1}{2} \int_0^T \text{Re} \left\{ \int_0^t [\exp(j2\pi\Delta f(t-s)) \right. \\ & \quad \left. - \exp(j2\pi\Delta f(s-t))] \right. \\ & \quad \left. \cdot h_4(t, s) r(s) ds r^*(t) dt \right\}. \end{aligned} \quad (39)$$

The receiver structure which implements (39) is known as an estimator-correlator receiver and is discussed in detail in Kailath [19]. This surprisingly straightforward structure is further simplified by the development of a linear estimator as an approximation to the LSE. Note that this receiver depends only on the single parameter w in $h_4(t, s)$ and was derived for small phase noise. When the phase noise is not small it is difficult to pursue alternative receiver types but quite simple to keep the structure and optimize the value of w . Hence the results for FSK are given for optimal w values. Simulations of the receiver in (39) give the BER results shown in Fig. 1.

For OOK systems, we have

$$\begin{aligned} H_1: \text{Data} = 1 \\ H_2: \text{Data} = 0. \end{aligned} \quad (40)$$

This corresponds to the following hypotheses using the notation in (20) and (21):

$$\begin{aligned} H_1: \alpha_1 &= 0, S_1 = S \\ H_2: \alpha_2 &= 0, S_2 = 0. \end{aligned} \quad (41)$$

Substituting these values in (35) gives

$$\begin{aligned} \hat{z}_1(t) &= \frac{1}{2} \int_0^t h_4(t, s) r(s) ds \\ \hat{z}_2(t) &= 0. \end{aligned} \quad (42)$$

Hence, from (19) the improved receiver has the form

$$\begin{aligned} \text{MLR} &= \int_0^T \text{Re} \left\{ \frac{1}{2} \int_0^t h_4(t, s) r(s) ds r^*(t) \right\} dt \\ &\quad - \frac{1}{2} \int_0^T \left| \frac{1}{2} \int_0^t h_4(t, s) r(s) ds \right|^2 dt. \end{aligned} \quad (43)$$

IV. IMPROVED RECEIVER USING THE OPTIMAL LINEAR LSE

A. The Optimal Linear LSE

As discussed above, it is very difficult in general to find $\hat{z}(t)$. However, $\hat{z}(t)$ is known to be a linear functional of past $r(\tau)$ determined by (11)–(14) when the signal process $z(t)$ is Gaussian. In this section we use the optimal linear functional of past $r(\tau)$ instead of the unknown least squares estimator $\hat{z}(t)$, to give alternative improved receivers. This must result in an improvement over Section III since the linear LSE derived previously was from an approximate Gaussian assumption rather than an optimization over all linear functionals.

We still drop the subscripts and the dependence on H_i for convenience and consider the LSE $\hat{z}(t)$ of the general $z(t)$ process in (21). First, we set

$$X(t) \stackrel{d.f.}{=} \alpha t + \theta(t) + \phi \quad (44)$$

$$\Delta(t, \tau) \stackrel{d.f.}{=} \begin{bmatrix} \cos(\alpha(t-\tau)) & -\sin(\alpha(t-\tau)) \\ \sin(\alpha(t-\tau)) & \cos(\alpha(t-\tau)) \end{bmatrix} \quad (45)$$

and observe that

$$\Delta^{-1}(t, \tau) = \begin{bmatrix} \cos(\alpha(t-\tau)) & \sin(\alpha(t-\tau)) \\ -\sin(\alpha(t-\tau)) & \cos(\alpha(t-\tau)) \end{bmatrix} \quad (46)$$

$$\Delta(s, \tau) \Delta^{-1}(t, \tau) = \Delta^{-1}(t, s). \quad (47)$$

Since

$$\begin{aligned} \cos X(t) \cos X(s) &= \frac{1}{2} [\cos(X(t) + X(s)) \\ &\quad + \cos(X(t) - X(s))] \end{aligned}$$

and

$$\begin{aligned} E[\cos(X(t) + X(s))] &= \text{Re}\{E[\exp(j(X(t) + X(s)))]\} \\ &= \text{Re}\{E[\exp(2j\phi)]E[\exp(j(\theta(t) + \theta(s) + \alpha(t+s)))]\} \\ &= 0 \end{aligned}$$

$$\begin{aligned} E[\cos(X(t) - X(s))] &= \text{Re}\{E[\exp(j(X(t) - X(s)))]\} \\ &= \text{Re}\{E[\exp(j(\theta(t) - \theta(s) + \alpha(t-s)))]\} \\ &= \cos(\alpha(t-s)) \exp(-\pi\beta|t-s|) \end{aligned}$$

we have

$$E[\cos X(t) \cos X(s)] = \frac{1}{2} \cos(\alpha(t-s)) \exp(-\pi\beta|t-s|). \quad (48)$$

Similarly

$$E[\cos X(t) \sin X(s)] = -\frac{1}{2} \sin(\alpha(t-s)) \exp(-\pi\beta|t-s|) \quad (49)$$

$$E[\sin X(t) \cos X(s)] = \frac{1}{2} \sin(\alpha(t-s)) \exp(-\pi\beta|t-s|) \quad (50)$$

$$E[\sin X(t) \sin X(s)] = \frac{1}{2} \cos(\alpha(t-s)) \exp(-\pi\beta|t-s|). \quad (51)$$

Using (48)–(51) and $E[z(t)] = 0$, $K(t, s)$ can be calculated by a full description of $z(t)$ rather than the approximating Gaussian

$$\begin{aligned} K(t, s) &= E[z(t)z^T(s)] \\ &= E \left[S^2 \begin{pmatrix} \cos X(t) \\ \sin X(t) \end{pmatrix} \begin{pmatrix} \cos X(s) & \sin X(s) \end{pmatrix} \right] \\ &= S^2 E \begin{bmatrix} \cos X(t) \cos X(s) & \cos X(t) \sin X(s) \\ \sin X(t) \cos X(s) & \sin X(t) \sin X(s) \end{bmatrix} \\ &= \frac{1}{2} S^2 \exp(-\pi\beta|t-s|) \\ &\quad \cdot \begin{bmatrix} \cos(\alpha(t-s)) & -\sin(\alpha(t-s)) \\ \sin(\alpha(t-s)) & \cos(\alpha(t-s)) \end{bmatrix}. \end{aligned}$$

Therefore

$$K(t, s) = \frac{1}{2} S^2 \exp(-\pi\beta|t-s|) \Delta(t, s). \quad (52)$$

Using (52) in the Wiener–Hopf equation, we have for $0 \leq s < t$

$$\begin{aligned} h(t, s) + \frac{1}{2} S^2 \int_0^t h(t, \tau) \exp(-\pi\beta|\tau-s|) \Delta(\tau, s) d\tau \\ = \frac{1}{2} S^2 \exp(-\pi\beta(t-s)) \Delta(t, s). \end{aligned}$$

Hence

$$\begin{aligned} h(t, s) \Delta^{-1}(t, s) + \frac{1}{2} S^2 \int_0^t h(t, \tau) \\ \cdot \exp(-\pi\beta|\tau-s|) \Delta(\tau, s) \Delta^{-1}(t, s) d\tau \\ = \frac{1}{2} S^2 \exp(-\pi\beta(t-s)) I. \end{aligned}$$

From (47), the above equation becomes

$$\begin{aligned} h(t, s) \Delta^{-1}(t, s) + \frac{1}{2} S^2 \int_0^t h(t, \tau) \\ \cdot \exp(-\pi\beta|\tau-s|) \Delta^{-1}(t, \tau) d\tau \\ = \frac{1}{2} S^2 \exp(-\pi\beta(t-s)) I. \end{aligned}$$

Let

$$H(t, s) \stackrel{d.f.}{=} h(t, s) \Delta^{-1}(t, s) \quad (53)$$

then we have

$$\begin{aligned} H(t, s) + \frac{1}{2}S^2 \int_0^t H(t, \tau) \exp(-\pi\beta|\tau - s|) d\tau \\ = \frac{1}{2}S^2 \exp(-\pi\beta(t - s))I. \end{aligned} \quad (54)$$

We can show $H(t, s)$ has the following form since the right-hand side of (54) is a diagonal matrix with identical entries

$$H(t, s) = \begin{pmatrix} H_{11}(t, s) & 0 \\ 0 & H_{11}(t, s) \end{pmatrix}. \quad (55)$$

Also $H_{11}(t, s)$ satisfies the equation

$$\begin{aligned} H_{11}(t, s) + \frac{1}{2}S^2 \int_0^t H_{11}(t, \tau) \exp(-\pi\beta|\tau - s|) d\tau \\ = \frac{1}{2}S^2 \exp(-\pi\beta(t - s)). \end{aligned} \quad (56)$$

After converting (56) to a differential equation (see Appendix A) the solution of (56) can be derived as

$$H_{11}(t, s) = \frac{(b - \pi\beta)[a \exp(bs) + \exp(-bs)]}{a \exp(bt) - a^{-1} \exp(-bt)} \quad (57)$$

$$b = \sqrt{\pi\beta(S^2 + \pi\beta)}, \quad a = \frac{b + \pi\beta}{b - \pi\beta}. \quad (58)$$

From (53) to (58), we get the solution of the Wiener-Hopf equation

$$h(t, s) = H_{11}(t, s)\Delta(t, s) \quad (59)$$

and using (13) after a little algebra we give a linear approximation to $\hat{z}(t)$ as

$$\hat{z}(t) = \int_0^t H_{11}(t, s) \exp(j\alpha(t - s))r(s) ds. \quad (60)$$

Note that using the exact covariance structure in $K(t, s)$ yields the optimal linear functional whereas the small-phase approach is only optimal for Gaussian signals.

B. Improved Receivers

For FSK systems the improved receiver is given by

$$\begin{aligned} \text{MLR} = \int_0^T \text{Re} \left\{ \int_0^t H_{11}(t, s) \right. \\ \cdot [\exp(j2\pi\Delta f(t - s)) - \exp(j2\pi\Delta f(s - t))] \\ \cdot r(s) ds r^*(t) \Big\} dt. \end{aligned} \quad (61)$$

Note that the receiver structure suggested by (61) is equivalent to the estimator-correlator structure discussed for (39), but with $H_{11}(t, s)$ replacing $h_4(t, s)$. Simulated BER results are shown for this receiver in Fig. 2.

For OOK systems we have the improved receiver

$$\begin{aligned} \text{MLR} = \int_0^T \text{Re} \left\{ \frac{1}{2} \int_0^t H_{11}(t, s) r(s) ds r^*(t) \right\} dt \\ - \frac{1}{2} \int_0^T \left| \int_0^t H_{11}(t, s) r(s) ds \right|^2 dt. \end{aligned} \quad (62)$$

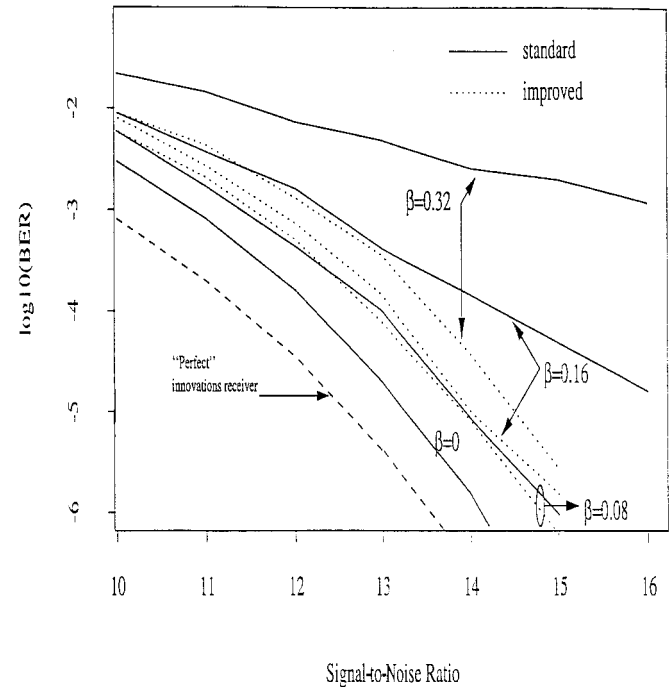


Fig. 2. Improved receiver using the optimal linear LSE.

V. CONCLUSIONS

To evaluate the improved FSK receivers in Sections III and IV, we compare their performance to the standard receiver discussed by Foschini *et al.* [2] where the IF input is passed through two branches each containing IF filters (at different center frequencies) and square law envelope detectors. The branches are then differenced and the output is compared to a detection threshold of zero. We also compare results to two idealized measures of performance, the zero phase noise case ($\beta = 0$) and a "perfect" innovations receiver ($\hat{z}(\cdot) = z(\cdot)$). The BER of both ideal cases is straightforward if a little awkward to compute and details can be found in [25]. Results are presented for a variety of normalized linewidths $\zeta = \beta T$ where we set $T = 1$ without loss of generality and have $\zeta = \beta$.

The improved receivers are analyzed by simulation in the following way. For each replicate a Brownian motion path is simulated for the phase noise process which gives a discretized version of $z(t)$ in $[0, T]$ from (21). The MLR is then computed from (39) or (61) which depends only on constants, the functions $h_4(\cdot)$ (30), $H_{11}(\cdot)$ (57), and $r(t) = z(t) + n(t)$. With $z(t)$ simulated as above the remaining term is the white noise process $n(t)$ which gives the stochastic integrals whose computation is discussed in Section II and in [20]. The MLR is then compared to a detection threshold of zero and the whole process is replicated to achieved reasonable BER estimates.

From Figs. 1 and 2 we can see that the improved receivers derived from both the small phase approach and the optimal linear approach do improve performance markedly at $\beta = 0.32, 0.16$. For $\beta = 0.08$ there is very little difference in performance, but then for small β we are nearing the region where the standard receiver is optimal. The best linear approach offers some performance over the small phase approach as you would expect as the small phase approach is a special

case of the linear approach. Both improved receivers have the simple estimator-correlator structure and intuition discussed by Kailath [20], [21]. In addition the functions $h_4(t, s)$ in (30) and $H_{11}(t, s)$ in (61) and (62) can be thought of as weighting functions, as in Kaiser *et al.* [12], [13], [17], to counter the phase noise wander which increases in variability over the bit period.

Hence, this work shows the degree of performance gain that can be achieved with simple receiver structures implementing an intuitive “weighting” of the phase noise wander. Furthermore, the results cast some light on the difficulty of approaching the $\beta = 0$ limit suggesting that true optimal performance lies some way above this bound. The tantalizing question in this area is how much improvement is possible. The fact that an optimal structure using an optimal linear estimator to the LSE does not approach $\beta = 0$ adds further to the mounting evidence that huge performance gains are not likely to be possible irrespective of receiver complexity. Note that at low signal-to-noise ratio (SNR), low β values the standard receiver can outperform both improved receivers. The reason is that in this region the errors are primarily caused by the white noise process and the standard receiver is optimal in the white noise only case.

APPENDIX A THE SOLUTION OF (56)

We assume that the solution of (56) has the form

$$H_{11}(t, s) = f(t)g(s) \quad (63)$$

where $f(t)$ and $g(s)$ are functions with second-order derivatives. Substituting $f(t)g(s)$ for $H_{11}(t, s)$ in (56), we get

$$\begin{aligned} f(t)g(s) + \frac{1}{2}S^2 \int_0^t f(t)g(\tau) \exp(-\pi\beta|\tau - s|) d\tau \\ = \frac{1}{2}S^2 \exp(-\pi\beta(t - s)). \end{aligned} \quad (64)$$

Rearranging gives

$$\begin{aligned} f(t) \left[2S^{-2}g(s) + \int_0^s g(\tau) \exp(-\pi\beta(s - \tau)) d\tau \right. \\ \left. + \int_s^t g(\tau) \exp(-\pi\beta(\tau - s)) d\tau \right] \\ = \exp(-\pi\beta(t - s)). \end{aligned} \quad (65)$$

Differentiating both sides of the above equation with respect to the variable s gives

$$\begin{aligned} f(t) \left\{ 2S^{-2}g^{(1)}(s) + \pi\beta \left[- \int_0^s g(\tau) \exp(\pi\beta(\tau - s)) d\tau \right. \right. \\ \left. \left. + \int_s^t g(\tau) \exp(-\pi\beta(\tau - s)) d\tau \right] \right\} \\ = \pi\beta \exp(-\pi\beta(t - s)). \end{aligned}$$

Hence

$$\begin{aligned} f(t) \left\{ \pi\beta \left[2S^{-2}g(s) + \int_0^s g(\tau) \exp(\pi\beta(\tau - s)) d\tau \right. \right. \\ \left. \left. + \int_s^t g(\tau) \exp(-\pi\beta(\tau - s)) d\tau \right] \right. \\ \left. + \left[2S^{-2}g^{(1)}(s) - 2S^{-2}\pi\beta g(s) - 2\pi\beta \int_0^s g(\tau) \right. \right. \\ \left. \left. \cdot \exp(\pi\beta(\tau - s)) d\tau \right] \right\} \\ = \pi\beta \exp(-\pi\beta(t - s)). \end{aligned}$$

Substituting (65) into the above gives

$$\begin{aligned} f(t) \left[g^{(1)}(s) - \pi\beta g(s) - \pi\beta S^2 \int_0^s g(\tau) \exp(\pi\beta(\tau - s)) d\tau \right] \\ = 0. \end{aligned}$$

If $f(t) \neq 0$, the above equation becomes

$$g^{(1)}(s) - \pi\beta g(s) - \pi\beta S^2 \int_0^s g(\tau) \exp(\pi\beta(\tau - s)) d\tau = 0. \quad (66)$$

Differentiating both sides of (66) we have

$$\begin{aligned} g^{(2)}(s) - \pi\beta \left[g^{(1)}(s) - \pi\beta S^2 \int_0^s g(\tau) \right. \\ \left. \cdot \exp(\pi\beta(\tau - s)) d\tau + S^2 g(s) \right] = 0. \end{aligned}$$

From (66) the above equation becomes

$$g^{(2)}(s) - \pi\beta(S^2 + \pi\beta)g(s) = 0. \quad (67)$$

This is a standard second-order differential equation with constant coefficients. The general solution of (67) is given by

$$g(s) = C_1 \exp(bs) + C_2 \exp(-bs) \quad (68)$$

where C_1 and C_2 are constants and $b = \sqrt{\pi\beta(S^2 + \pi\beta)}$. From (66), we know

$$g^{(1)}(0) - \pi\beta g(0) = 0.$$

Hence, using (68), we have

$$b(C_1 - C_2) - \pi\beta(C_1 + C_2) = 0$$

and

$$C_1 = \frac{b + \pi\beta}{b - \pi\beta} C_2.$$

We choose $C_2 = 1$ and then get

$$g(s) = a \exp(bs) + \exp(-bs) \quad (69)$$

where

$$a = \frac{b + \pi\beta}{b - \pi\beta}.$$

Substituting (69) into (65) gives after some messy calculations

$$f(t) = \frac{b - \pi\beta}{a \exp(bt) - a^{-1} \exp(-bt)}. \quad (70)$$

Therefore, from (63)

$$H_{11}(t, s) = \frac{(b - \pi\beta)(a \exp(bs) + \exp(-bs))}{a \exp(bt) - a^{-1} \exp(-bt)} \quad (71)$$

where a and b are given by (58). Thus, we have shown that (57) and (58) is the solution of (56).

REFERENCES

- [1] G. J. Foschini and G. Vannucci, "Characterising filtered lightwaves corrupted by phase noise," *IEEE Trans. Inform. Theory*, vol. 34, pp. 1437–1448, Nov. 1988.
- [2] G. J. Foschini and L. J. Greenstein, "Noncoherent detection of coherent lightwave signals corrupted by phase noise," *IEEE Trans. Commun.*, vol. 36, pp. 306–314, Mar. 1988.
- [3] I. Garrett and G. Jacobsen, "Influence of (semiconductor) laser linewidth on the error-rate floor in dual filter optical FSK receivers," *Electron. Lett.*, vol. 21, pp. 280–282, 1985.
- [4] ———, "Theoretical analysis of heterodyne optical receivers for transmission systems using (semiconductor) lasers with nonnegligible linewidth," *J. Lightwave Technol.*, vol. LT-4, pp. 323–334, 1986.
- [5] ———, "The effect of laser linewidth on coherent optical receivers with nonsynchronous demodulation," *J. Lightwave Technol.*, vol. LT-5, pp. 551–560, Apr. 1987.
- [6] ———, "Theory for optical heterodyne DPSK receivers with post-detection filtering," *J. Lightwave Technol.*, vol. LT-5, pp. 478–484, Apr. 1987.
- [7] G. Jacobsen and I. Garrett, "Theory for optical heterodyne ASK receivers using square-law detection and post-detection filtering," *Proc. IEEE*, vol. 134, pp. 302–312, 1987 and see also erratum, vol. 135, p. 100, 1988.
- [8] I. Garrett and G. Jacobsen, "Possibilities for coherent optical communication systems using lasers with large phase noise," *Brit. Telecom Technol. J.*, vol. 7, no. 4, pp. 5–11, Oct. 1989.
- [9] I. Garrett, K. J. Bond, J. B. Waite, D. S. L. Lettis, and G. Jacobsen, "Impact of phase noise in weakly coherent systems: A new and accurate approach," *J. Lightwave Technol.*, vol. 8, pp. 329–337, Mar. 1990.
- [10] G. Jacobsen and I. Garrett, "The effect of crosstalk and phase noise in multichannel coherent optical ASK systems," *J. Lightwave Technol.*, vol. 9, pp. 1006–1018, 1991.
- [11] Y. E. Dallah and S. Shamai (Shitz), "Time diversity in DPSK noisy phase channels," *IEEE Trans. Commun.*, vol. 40, pp. 1703–1715, Nov. 1992.
- [12] C. P. Kaiser, P. J. Smith, and M. Shafi, "An improved optical heterodyne DPSK receiver to combat laser phase noise," *J. Lightwave Technol.*, vol. 13, pp. 525–533, Mar. 1995.
- [13] C. P. Kaiser, "Optical heterodyne binary-DPSK systems," Ph.D. dissertation, Dep. Elec. Electron. Eng., Univ. Canterbury, Christchurch, New Zealand, Dec. 1993.
- [14] J. Salz, "Coherent lightwave communications," *AT&T Tech. J.*, vol. 64, pp. 2153–2209, Dec. 1985.
- [15] D. J. Bond, "The statistical properties of phase noise," *Brit. Telecom Technol. J.*, vol. 7, pp. 12–17, Oct. 1989.
- [16] G. L. Pierobon and L. Tomba, "Moment characterization of phase noise in coherent optical systems," *J. Lightwave Technol.*, vol. 9, pp. 996–1005, Aug. 1991.
- [17] C. P. Kaiser, M. Shafi, and P. J. Smith, "Analysis methods for optical heterodyne DPSK receivers corrupted by laser phase noise," *J. Lightwave Technol.*, vol. 11, pp. 1820–1830, Nov. 1993.
- [18] T. Kailath, "The innovations approach to detection and estimation theory," *Proc. IEEE*, vol. 58, pp. 680–690, May 1970.
- [19] ———, "A general likelihood-ratio formula for random signals in Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 350–361, May 1969.
- [20] ———, "Likelihood ratios for Gaussian processes," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 276–288, May 1970.
- [21] ———, "A further note on a general likelihood formula for random signals in Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 393–396, July 1970.
- [22] C. N. Georghiades, "A proposed receiver structure for heterodyne optical communication systems," Ph.D. dissertation, Dep. Elec. Eng., Washington Univ., St. Louis, MO, May 1985.
- [23] C. N. Georghiades and K. L. Snyder, "Receiver performance for heterodyne optical communication," *IEEE Trans. Commun.*, vol. COM-34, pp. 1096–1104, Nov. 1986.
- [24] P. J. Smith, M. Shafi, and H. Gao, "Optical heterodyne binary-DPSK systems: A review of analysis and performance," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 557–568, Apr. 1995.
- [25] H. Gao, "The design and analysis of some computationally intensive communication systems," Ph.D. dissertation, Inst. Stat. Operations Res., Victoria Univ. Wellington, Wellington, New Zealand, June 1996.

Hongsheng Gao received the B.Sc. and M.Sc. degrees in mathematics from Jilin University, China, in 1982 and 1985 respectively. Since 1993, he has been at Victoria University, Wellington, New Zealand, where he has recently received the Ph.D. degree.

From 1985 to 1993 he worked for the Changchun Post and Telecommunications Institute. His research interests are in pure algebra, ring theory and statistics, especially the statistical analysis of antenna arrays.

Peter J. Smith (M'93) received the B.Sc. degree in mathematics and the Ph.D. degree in statistics from the University of London, London, U.K., in 1983 and 1988, respectively.

From 1983 to 1986, he was with the Telecommunications Laboratories at GEC Hirst Research Centre. Since 1988, he has been a lecturer in statistics at Victoria University, Wellington, New Zealand. His research interests include the statistical aspects of design and analysis for communication systems, especially coherent optical systems, antenna arrays, and mobile radio.

Mansoor Shafi (S'69–A'70–M'82–SM'87–F'93) received the B.S. degree from the Engineering University, Lahore, Pakistan, and the Ph.D. degree from the University of Auckland, Auckland, New Zealand, in 1970 and 1979, respectively, both in electrical engineering.

Since 1979, he has held different positions with the New Zealand Post Office and is currently a Wireless Architecture Manager with Telecom New Zealand. His research interests are in fixed and mobile transmission systems.

Dr. Shafi was awarded the IEEE Communications Society Public Service Award in Communications in 1992.