Adaptive Channel Estimator for an HF Radio Link

A. P. CLARK AND S. HARIHARAN

Abstract—**Several channel estimators have been developed over the past few years for use in serial digital modems that operate over voiceband HF radio links. A simple estimator designed for a 2400 bit/s modem is a development of the conventional gradient estimator, and employs a polynomial filter that gives a prediction of the channel response. An improved estimator designed for a 9600 bit/s modem is much more sophisticated, and uses the techniques of the simple estimator together with a knowledge of the number of different paths (separate sky waves) in the HF radio link. The paper describes some new estimators, that are developments of the simple estimator but make no use of any knowledge of the number of different paths. The new estimators have performances intermediate between those of the simple and improved estimators, but are only a little more complex than the simple estimator. The new estimators are studied for a particular QPSK modem that operates at 4800 bits/s over voiceband HF radio links. Results of computer simulation tests are presented, comparing the accuracies of the channel estimates given by different estimators, and hence suggesting the most suitable of these.**

I. INTRODUCTION

THE paper is concerned with the design of a channel estimator for a digital data-transmission system, operating with a serial 4800 bit/s signal over a voiceband HF radio link, as modeled in Fig. 1. The signal fed to the voiceband channel is a quarternary phase-shift-keyed (QPSK) signal with a carrier frequency of 1800 Hz and a signal element rate of 2400 bauds. Linear modulation and demodulation processes are used, and the receiver forms an estimate of the sampled impulse-response of the resultant baseband channel, formed by the combination of the low-pass filter, modulator, transmission path, and demodulator (Fig. 1).

Considerable advances have been achieved over the past few years in the design of serial modems for HF radio links [1]–[11]. As a result of this the highest practically obtainable transmission rate over a voiceband HF channel has been increased from 2400 to 9600 bits/s [1]–[11]. The increase has been achieved through the development of more effective techniques for tracking the sampled impulse-response of the time-varying baseband channel involving the HF radio link [12]–[29], together with the development of more effective detection processes for handling the severe signal distortion introduced by the HF radio link [30]–[33]. Considerable attention has been paid to the Kalman filter as a means for holding the receiver correctly adjusted for the channel [12]–[25]. A Kalman filter assumes that the channel performs a degree-1 Markov process on the signal [16], which is a valid assumption for both time invariant and random-walk channels [16]. Thus, a Kalman filter is optimum for either of these two channels, in the sense that it can give the minimum mean-square error in the adaptive adjustment of the receiver. Unfortunately, a typical HF radio link cannot be modeled as a degree-1 Markov process, and computer-simulation tests have, in fact, confirmed that the conventional Kalman filter [12]–[19], together with its more recent developments [20]–[25] are not optimum for a typical HF channel [25], but could be made so through a further considerable increase in equipment complexity [20], [25]. In particular, the use of a window with a Kalman filter [24] does not compensate for the fact that the Kalman filter itself is based on the incorrect assumption that the channel is time invariant [25]. A further weakness of the Kalman filter is the considerable equipment complexity involved.

In view of these considerations, an alternative approach has been studied in which the correct adjustment of the receiver requires a knowledge only of the sampled impulse-response of the channel [26]–[29]. When an adaptive filter is used ahead of the detector, the filter is now adjusted directly from an estimate of the channel sampled impulse-response [32], rather than, for example, by minimizing the mean-square error in the equalized signal [15], [17], [19], [23]. It is, in general, much easier to estimate the sampled impulse-response of the channel than to adjust a filter that is some function of the inverse of the channel response, as in a conventional adaptive equalizer [10], [25]. As a result of the investigations, some promising channel estimators have been developed [27]–[29]. Computer-simulation tests over various models of an HF radio link have shown that, with the most effective of the estimators [29], satisfactory operation is obtained at 9600 bits/s under quite severe fading conditions [9], [33]. The estimator here, although complex, is not unduly so and has a much better performance than the corresponding Kalman-filter estimator [25], [29]. However, the estimator requires a knowledge of the number of different paths (separate sky waves) in the HF radio link, and its correct operation is rather critically dependent on the use of an adequate starting-up procedure [34]. It is also considerably more complex than the simple estimator that operates well at 2400 bits/s [1], [28]. The latter estimator is, in fact, only slightly more complex than the conventional gradient estimator [26], [27], but has a very much better performance [28]. At a transmission rate of 4800 bits/s, the accuracy needed in the channel estimate is intermediate between those for 2400 and 9600 bits/s [28], [29]. This suggests that the best channel estimator [29], is perhaps needlessly sophisticated, so that satisfactory operation at 4800 bits/s could well be achieved with a significantly less complex estimator. Since the main source of complexity in the estimator is involved with the modeling of the multipath propagation in the HF radio link, it was felt that the best approach towards a simpler but adequate estimator would be to develop the simple estimator [28], but without requiring this to model the multipath propagation, that is, the number of different paths, their relative transmission delays, and so on.

The paper describes various developments of the simple estimator, and then presents the results of computer-simulation tests on the resulting systems. To enable the tests to be carried out within a reasonable period of time and to enable reasonably precise comparisons to be made between different systems, some idealistic assumptions have been made concerning the fading and noise introduced by the HF radio link. The tests therefore represent an initial feasibility study, whose aim

Paper approved by the Editor for Radio Communications of the IEEE Communications Society. Manuscript received September 21, 1987; revised August 1, 1988.

The authors are with the Department of Electronic and Electrical Engineering, Loughborough University of Technology, Loughborough, Leics., LE11 3TU, England.

IEEE Log Number 8929591.

0090-6778/89/0900-0918$01.00

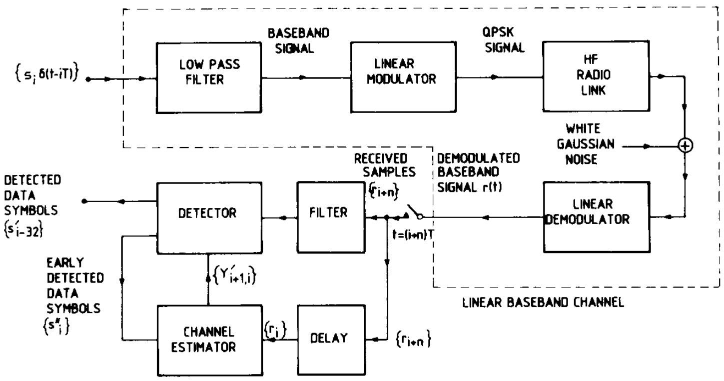


Fig. 1. Model of data-transmission system.

is to select the more promising systems for further and more realistic tests.

II. Model of System

The model of the data-transmission system is shown in Fig. 1. The data symbols $\{s_i\}$ are statistically independent and equally likely to have any of their four possible values, which are given by

$$s_i = s_{0,i} + js_{1,i} \tag{1}$$

where $j = \sqrt{-1}$, and

$$s_{0,i},\ s_{1,i} = \pm 1. \tag{2}$$

A stream of signal elements, in the form of the regularly spaced impulses $\{s_i\delta(t - iT)\}$, is fed to the low-pass filter, whose output complex-valued continuous waveform is the baseband modulating waveform that is fed to the linear modulator. The real and imaginary parts of a complex-valued signal are, of course, separate signals, so that two separate but identical lowpass filters are in fact employed here, each feeding the corresponding output signal to a separate input terminal of the linear modulator. The output signal from the linear modulator is a serial stream of real-valued QPSK signal elements, with a carrier frequency of 1800 Hz and an element rate of 2400 bauds. Each signal element itself comprises the sum of two binary double sideband suppressed carrier amplitude modulated elements, with their carriers in phase quadrature, the binary values of the in-phase and quadrature elements being determined, respectively, by the real and imaginary parts ($s_{0,i}$ and $s_{1,i}$) of the corresponding data-symbol s_i. Thus, the QPSK signal is handled as a quadrature amplitude modulated (QAM) signal.

The QPSK signal is fed to the HF radio link, where its spectrum is shifted into the HF band, this signal being then transmitted via three different Rayleigh-fading paths (three independent sky waves) to the receiver, where its spectrum is returned to the voice band. The transmission delays of the three paths are assumed to be constant. Stationary white Gaussian noise, with zero mean and a two-sided power spectral density of $(1/2)N_0$, is added to the data signal at the output of the HF radio link.

Each of the three Rayleigh-fading paths is modeled theoretically as shown in Fig. 2. The waveforms $f_1(t)$ and $f_2(t)$ here are real-valued narrow-band baseband Gaussian waveforms, whose spectral shaping is approximately Gaussian. Each of the waveforms $f_1(t)$ and $f_2(t)$ is formed by feeding white Gaussian noise through an appropriate five-pole Bessel filter, which gives a good approximation to the ideal Gaussian spectral shaping over the effective bandwidth [28].

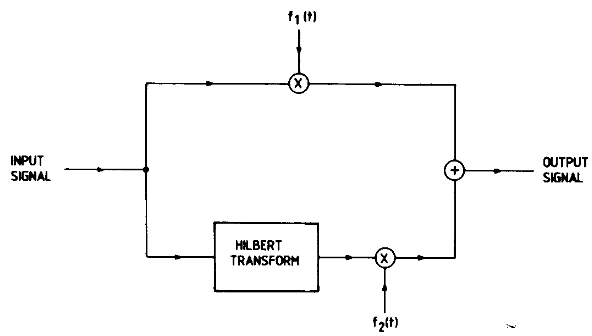


Fig. 2. Rayleigh fading introduced by a sky wave.

The six Gaussian waveforms involved in the three fading paths are statistically independent with zero mean, the same variance and the same root mean square bandwidth which is 1 Hz in every case. Thus, the signal received over each path has the same mean-square value and the same frequency spread of 2 Hz. The transmission delays of the three paths, measured relative to that of the first path, are 0, 1.1, and 3 ms. Three fading paths, a time spread of 3 ms and a frequency spread of 2 Hz have been used to provide a more severe test of the channel estimators than the model for poor conditions suggested by CCIR [35]. The channel model used here is, however, firmly based on the CCIR recommendations. The selected time delays ensure a different sampled impulse response for each path (sky wave) and avoid some rather exceptional effects that occur in the adjustment of the adaptive filter (Fig. 1), when one of the relative (differential) time delays is an integral multiple of the other. Details of these are beyond the scope of this paper. The average transmitted energy per bit of information, at both the input and output of the HF radio link is arranged to be unity, no change in the average signal level being therefore introduced by the HF link.

The linear demodulator in Fig. 1 filters and demodulates the received signal, using two linear coherent demodulators whose reference carriers are in-phase quadrature and have a constant frequency. This is equal to the average instantaneous frequency of the received signal carrier, thus eliminating any constant frequency offset in the received QPSK signal, but not tracking the variations in instantaneous frequency introduced by the HF radio link, the whole of which therefore appears in the demodulated waveform $r(t)$. The demodulated signals at the outputs of the in-phase and quadrature coherent demodulators are taken to be the real and imaginary parts, respectively, of a complex variable, so that the resultant demodulated baseband signal $r(t)$ is complex valued. The latter is sampled once per data symbol, and the sample at time $t = iT$ is

$$r_i = \sum_{h=0}^{g} s_{i-h} y_{i,h} + w_i$$

$$= Y_i S_i^T + w_i \tag{3}$$

where

$$Y_i = [y_{i,0}\ y_{i,1}\ \cdots\ y_{i,g}] \tag{4}$$

and

$$S_i = [s_i\ s_{i-1}\ \cdots\ s_{i-g}]. \tag{5}$$

Y_i and S_i are $(g + 1)$-component row vectors, and S_i^T is the transpose of S_i. The vector Y_i is taken to be the sampled impulse-response of the linear baseband channel in Fig. 1. This channel is formed by the low-pass filter, linear modulator, HF radio link and linear demodulator. The scalar quantity w_i is a noise component originating from the white Gaussian noise (Fig. 1). The $\{r_i\}$, $\{y_{i,h}\}$, and $\{w_i\}$ are all complex valued. The filters in the linear demodulator are such that the real and imaginary parts of the $\{w_i\}$ are Gaussian random variables with zero mean and variance that is dependent on $(1/2)N_0$, neighboring $\{w_i\}$ being slightly correlated [9], [31]. The receiver input filters are representative of those in a practical modem, whose input filters are connected in cascade with typical filters used in the radio receiving equipment. Further details of the filters used in the transmitter and receiver of the modem are given elsewhere [9], [31].

The received samples $\{r_i\}$ are fed to an adaptive linear feedforward transversal filter. The latter is an all-pass network that adjusts the sampled impulse-response of the channel and filter to be minimum phase, without changing any amplitude distortion in the received signal [32]. With the aid of the adaptive filter, a near-optimum tolerance to noise can be achieved by means of a relatively simple detector, leading to a potentially cost-effective system [33].

The received samples $\{r_i\}$ are also fed to the channel estimator, after being suitably delayed. The channel estimator uses the received samples $r_{i-g}, r_{i-g+1}, \cdots, r_i$ together with the "early" detected data-symbols $s''_{i-g}, s''_{i-g+1}, \cdots, s''_i$ and the one-step prediction of Y_i, given by

$$Y'_{i,i-1} = [y'_{i,i-1,0}\ y'_{i,i-1,1}\ \cdots\ y'_{i,i-1,g}] \tag{6}$$

to form the updated estimate of Y_i, given by

$$Y'_i = [y'_{i,0}\ y'_{i,1}\ \cdots\ y'_{i,g}]. \tag{7}$$

Thus, $Y'_{i,i-1}$ is the evaluation (informed guess) that the estimator makes of Y_i, between the receipt of r_{i-1} and r_i, and Y'_i is the evaluation that it makes of Y_i, following the receipt of r_i. The estimator next forms the one-step prediction of Y_{i+1}, given by $Y'_{i+1,i}$. The latter is fed to the detector, ready for the next detection process that gives s''_{i+1}, and so on. Clearly, any error in $Y'_{i+1,i}$ correspondingly degrades the detection of s_{i+1}.

The "early" detected data symbols have no delay in detection. This minimizes the period over which prediction must be carried out but increases somewhat the error rate in the $\{s''_i\}$. The detected data-symbols $\{s'_i\}$ at the output of the detector (Fig. 1) have a delay in detection of 32 sampling intervals, no significant reduction in error rate being achieved by any further increase in the delay in detection. It is assumed here that one step prediction is used for the detector (6), and each estimator is tested by measuring the mean-square error in the corresponding one-step prediction $Y'_{i,i-1}$.

The important advantage gained by using the adaptive filter in Fig. 1 is that it avoids the need for prediction over many sampling intervals, such as must be used in the absence of the filter [1], [9]. Prediction over many sampling intervals can increase considerably the error in the prediction [13], [28]. Further details of the adaptive filter and detector are beyond the scope of this paper and are given elsewhere [10], [32], [33].

Since we are concerned here with the operation of the channel estimator and not with the detector, the correct detection of all data symbols is assumed, even at low signal/noise ratios, so that

$$s''_i = s_i \tag{8}$$

for all $\{i\}$. Tests have indicated that the performance of the channel estimator is only likely to be significantly affected by errors in the $\{s''_i\}$ at the higher error rates (above 10^{-2}) [1], [9]. There is, however, a more fundamental reason for the given assumption. In any practical application of the system, the data signal is divided into separate blocks, each preceded by a training signal whose data-symbol values are known at the receiver. Under fading conditions, such as those tested here, most errors in detection occur during the deeper fades and generally in long bursts. Often, during an error burst, the channel estimate becomes significantly degraded, leading to more errors in the $\{s''_i\}$, which, in turn, further degrade the channel estimate, and so on, until there is a complete failure of the system. The error burst is now extended to the end of the block of data symbols, but the following training signal restores correct operation of the channel estimator, ready for the next block of data symbols. When more than a few errors have occurred in the $\{s'_i\}$, for any block of data symbols, the whole block of detected data-symbols $\{s'_i\}$ is normally rendered invalid and is rejected by the receiver. Furthermore, any large burst of errors in the $\{s''_i\}$ is usually accompanied by a substantial burst of errors in the $\{s'_i\}$. It follows that, for the most reliable operation of the system, the channel estimator must give the most accurate possible estimate (prediction) of the channel when the $\{s''_i\}$ are correct, since once an appreciable burst of errors has occurred in the $\{s''_i\}$, the chances are that the corresponding block of $\{s'_i\}$ are invalid, and no advantage is gained by improving the channel estimate under these conditions. Another more practical reason for assuming (8) is the very considerable time and cost involved in testing the detector and channel estimator together, bearing in mind that this involves also the adjustment of the adaptive filter (Fig. 1) and the transmission of the appropriate training signals. The aim of the investigation described here is to select the most promising estimation process for further study, and the preferred system is, in fact, currently undergoing the more rigorous tests just mentioned. Details of these tests are however well beyond the scope of this paper.

III. System 1

Before describing the new channel estimator, it is necessary first to consider in some detail the principles behind the operation of the simple estimator [28]. The prediction process carried out by the estimator is, for convenience, considered first. The estimator uses the updated estimate of Y_i, given by Y'_i in (7), and the one-step prediction of Y_i, given by $Y'_{i,i-1}$ in (6), to determine an estimate of the error in the prediction, which is

$$X'_i = Y'_i - Y'_{i,i-1}. \tag{9}$$

The actual error in $Y'_{i,i-1}$ is, of course, $Y_i - Y'_{i,i-1}$.

The prediction of Y_{i+1} is now determined by means of a polynomial filter [13] that operates as follows:

$$Y''_{i+1,i} = Y''_{i,i-1} + (1-\theta)^2 X'_i \tag{10}$$

$$Y'_{i+1,i} = Y'_{i,i-1} + Y''_{i+1,i} + (1-\theta^2) X'_i. \tag{11}$$

The vector $Y'_{i+1,i}$ is the degree-1 least-squares fading-memory prediction of Y_{i+1} [13], [28], and the vector $Y''_{i+1,i}$ is a prediction of the rate of change with i of Y_{i+1}. These are considered in more detail elsewhere [13]. The symbol θ is a real-valued constant in the range 0 to 1, usually close to 1. At the start of the process,

$$Y''_{1,0}=0 \tag{12}$$

and

$$Y'_{1,0}=Y'_0 \tag{13}$$

where Y'_0 is determined from an appropriate synchronizing signal that precedes the transmission of data [34]. Extensive tests on the different versions of the prediction process in the simple estimator have shown that the above algorithm gives the smallest mean-square error in the prediction of Y_i, over the range of signal/noise ratios of greatest interest [1], [28].

To determine Y'_i in (9), the estimator forms an estimate r'_i of the received sample r_i, such that

$$r'_i=Y'_{i,i-1}S_i^T. \tag{14}$$

The vector S_i here [see (5)] is determined from the values of the corresponding $\{s''_{i-h}\}$, assuming that (8) holds. From (3), r'_i is the value of r_i when $Y'_{i,i-1} = Y_i$ and $w_i = 0$. The estimator next forms the error signal

$$e_i=r_i-r'_i \tag{15}$$

and it uses e_i to form the correction vector X'_i in (9), which is now given by

$$X'_i=be_iS_i^*. \tag{16}$$

The quantity b is an appropriate positive real-valued constant and S_i^* is the complex conjugate of S_i. The vector X'_i is added to $Y'_{i,i-1}$ to give the updated estimate

$$Y'_i=Y'_{i,i-1}+be_iS_i^* \tag{17}$$

from (9) and (16). Equation (17) is the conventional gradient or steepest descent algorithm for deriving the updated estimate of Y_i [26]–[28].

The channel estimators to be considered next are known here as Systems 2–6. These operate adaptively in such a way as to make a greater use of the available prior knowledge of Y_i than is achieved by system 1. The vector X'_i in (16) is now no longer constrained to lie in the direction of steepest descent, and it is adjusted adaptively to match the time-varying channel. The adaptive estimator does not however use any knowledge of the structure of the multipath channel formed by the HF radio link, that is, of the number of different paths (separate sky waves), their relative transmission delays, and so on.

IV. System 2

An estimate X'_i of the actual error in $Y'_{i,i-1}$, given by

$$X_i=Y_i-Y'_{i,i-1} \tag{18}$$

can, in principle, be derived from the fact that the prediction algorithm given by (9)–(11) employs a degree-1 least-squares fading memory polynomial filter [13], [28]. The latter assumes that the rate of change of Y_i with i is constant or only slowly varying with i. Thus, a significant source of error in a prediction $Y'_{i,i-1}$ is likely to be the acceleration (variation in rate of change) in Y_i. If the only error in $Y'_{i,i-1}$ is due to the acceleration in Y_i, then

$$Y_i=Y'_{i,i-1}+c_iA_i \tag{19}$$

where c_i is a complex-valued scalar and

$$\begin{aligned}A_i&=(Y_{i+1}-Y_i)-(Y_i-Y_{i-1})\\&=Y_{i+1}-2Y_i+Y_{i-1}\end{aligned} \tag{20}$$

such that $X_i = c_iA_i$, from (18). An estimate (prediction) of A_i is given by

$$A'_i=Y'_{i+1,i}-2Y'_{i,i-1}+Y'_{i-1,i-2}. \tag{21}$$

The weakness of A'_i in (21) is its relatively high noise level, bearing in mind that $Y'_{i-1,i-2}$, $Y'_{i,i-1}$, and $Y'_{i+1,i}$ do not differ greatly, much of the difference between them being due to the noise. Thus, instead of using A'_i, the estimator uses the corresponding vector

$$Z_i=[z_{i,0}\ \ z_{i,1}\ \ \cdots\ \ z_{i,g}] \tag{22}$$

which is derived from A'_i as follows. First, let the absolute value (modulus) of the $(h + 1)$th component of A'_i be $\alpha_{i,h}$, for $h = 0, 1, \cdots, g$, and suppose that A'_1 is the first of the $\{A'_i\}$ to be processed. Now $z_{i,h}$ is a measure of the average value of $\alpha_{i,h}$, which may be either the growing-memory average, given by

$$z_{i,h}=i^{-1}\sum_{l=1}^{i}\alpha_{l,h} \tag{23}$$

or else the fading-memory average, given by

$$z_{i,h}=a\sum_{l=1}^{i}(1-a)^{i-l}\alpha_{l,h} \tag{24}$$

where a is a real-valued constant such that $0 < a < 1$, and l is an integer. Equation (23) can be implemented sequentially as

$$\begin{aligned}z_{i,h}&=(1-i^{-1})z_{i-1,h}+i^{-1}\alpha_{i,h}\\&=z_{i-1,h}+i^{-1}(\alpha_{i,h}-z_{i-1,h})\end{aligned} \tag{25}$$

and (24) as

$$\begin{aligned}z_{i,h}&=(1-a)z_{i-1,h}+a\alpha_{i,h}\\&=z_{i-1,h}+a(\alpha_{i,h}-z_{i-1,h})\end{aligned} \tag{26}$$

where

$$z_{0,h}=0 \tag{27}$$

for $h = 0, 1, \cdots, g$.

Since all components of Z_i are real valued, whereas the components of A_i in (20) are, in general, complex valued, neither $Y'_{i,i-1} + Z_i$ nor $Y'_{i,i-1} + c_iZ_i$ could be used as a satisfactory updated estimate of Y_i in (19). Nevertheless, $z_{i,h}$ gives a measure of the magnitude $|y_{i,h} - y'_{i,i-1,h}|$ of the error in the component $y'_{i,i-1,h}$ of $Y'_{i,i-1}$. Furthermore, for the most accurate tracking of a time-varying channel, the step size employed in the gradient algorithm of (17) should be permitted to vary from one component of $Y'_{i,i-1}$ to another, and should increase with the likely magnitude of the error in that component. These considerations suggest that (17) should be replaced by

$$y'_{i,h}=y'_{i,i-1,h}+bu_{i,h}e_is^*_{i-h} \tag{28}$$

for $h = 0, 1, \cdots, g$ where b is an appropriate small positive real-valued constant, and

$$u_{i,h}=p(z_{i,h}). \tag{29}$$

$p(z_{i,h})$ is a monotonically nondecreasing positive real-valued function of $z_{i,h}$. The parameter $u_{i,h}$ in (28) cannot be replaced by $z_{i,h}$ itself, for the following reasons. First, no $u_{i,h}$ must be permitted to remain at zero for any significant period, since, if this occurs, the corresponding component of $Y''_{i,i-1}$ in (10) may become locked at zero, thus preventing any further change in the corresponding $y'_{i,i-1,h}$. Second, no $u_{i,h}$ should be permitted to become too large, in order to avoid possible

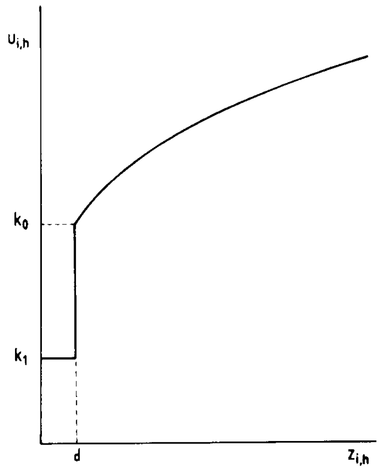


Fig. 3. Relationship between $u_{i,h}$ and $z_{i,h}$ for System 2.

instability of the algorithm given by (28). Thus, the value of $u_{i,h}$ should be constrained such that

$$k_1 < u_{i,h} < k_2 \tag{30}$$

where k_1 and k_2 are appropriate positive real-valued constants. Finally, tests have shown that, for the best performance, $u_{i,h}$ must vary nonlinearly with $z_{i,h}$ over the range k_1 to k_2. In the most effective arrangement that has been found, $u_{i,h}$ varies with $z_{i,h}$ as shown in Fig. 3 where $k_1 = 10^{-6}$ and $k_2 \to \infty$. The quantity k_o is a small positive real-valued constant,

$$d = k_o^4 \tag{31}$$

and when $z_{i,h} > d$,

$$u_{i,h} = z_{i,h}^{0.25}. \tag{32}$$

The nonlinear variation of $u_{i,h}$ with $z_{i,h}$ here prevents $u_{i,h}$ from becoming too large, so that it is not, in fact, necessary to limit the maximum value of $u_{i,h}$. As before, the prediction of Y_{i+1} is determined by (9), (10), and (11).

V. Systems 3–6

Instead of attempting to measure the acceleration in Y_i directly, use can be made of the fact that the greater the maximum magnitude of any $y_{i,h}$, the greater is likely to be its maximum acceleration and hence the greater the probable value of the largest error in the corresponding prediction $y'_{i,i-1,h}$. Systems 3–6 all operate on estimates of the magnitudes of the $\{y_{i,h}\}$, and, as before, the estimates are either growing-memory or fading-memory averages. In particular, the growing-memory average $x_{i,h}$ is now given by

$$x_{i,h}^2 = x_{i-1,h}^2 + i^{-1}(|y'_{i,i-1,h}|^2 - x_{i-1,h}^2) \tag{33}$$

and the fading-memory average $x_{i,h}$ is given by

$$x_{i,h}^2 = x_{i-1,h}^2 + a(|y'_{i,i-1,h}|^2 - x_{i-1,h}^2) \tag{34}$$

where a is a real-valued constant such that $0 < a < 1$, and

$$x_{1,h}^2 = |y'_{1,0,h}|^2 \tag{35}$$

for $h = 0, 1, \cdots, g$. (33) corresponds to (23) and (25), whereas (34) corresponds to (24) and (26). The gradient algorithm of (17) is here replaced by (28), with

$$u_{i,h} = p(x_{i,h}^2) \tag{36}$$

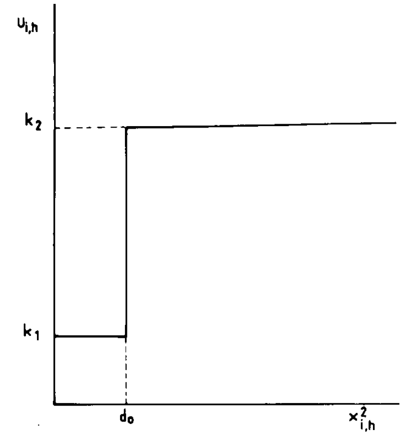


Fig. 4. Relationship between $u_{i,h}$ and $x_{i,h}^2$ for System 3.

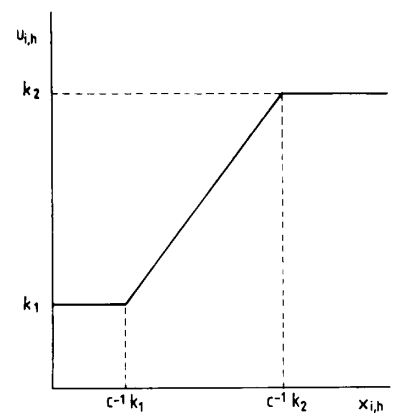


Fig. 5. Relationship between $u_{i,h}$ and $x_{i,h}$ for System 4.

and $p(\cdot)$ having the same basic properties as before. The prediction of Y_{i+1} is determined by (9)–(11), as before.

In System 3, $u_{i,h}$ varies with $x_{i,h}^2$ according to Fig. 4. Furthermore, with the particular HF radio link tested, the last ten components of Y_i are all ideally equal to zero. This leads to the nonadaptive version of System 3, in which the number of components of $Y'_{i,i-1}$ is reduced to 22, by simply setting to zero its last ten components and operating System 1 with the corresponding 22-component vector $Y'_{i,i-1}$. (17) is now used in place of (28) for the gradient algorithm.

System 4 is a simple modification of System 3, in which the relationship between $u_{i,h}$ and $x_{i,h}$ is as given by Fig. 5, such that

$$u_{i,h} = cx_{i,h} \tag{37}$$

when $k_1 < u_{i,h} < k_2$, and where c is an appropriate positive real-valued constant.

An interesting arrangement of System 4 is that where a in (34) is set to unity, so that

$$x_{i,h}^2 = |y'_{i,i-1,h}|^2 \tag{38}$$

and no averaging is in fact carried out.

In Systems 5 and 6, the relationships between $u_{i,h}$ and $x_{i,h}^2$ are as shown in Figs. 6 and 7, respectively. Over the curved

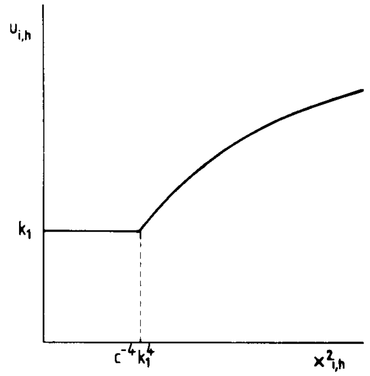


Fig. 6. Relationship between $u_{i,h}$ and $x^2_{i,h}$ for System 5.

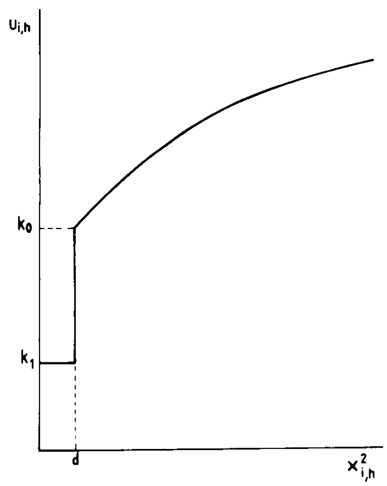


Fig. 7. Relationship between $u_{i,h}$ and $x^2_{i,h}$ for System 6.

portion of the relationship in Fig. 6,

$$u_{i,h}=cx^{0.5}_{i,h} \tag{39}$$

whereas, over the curved portion in Fig. 7,

$$u_{i,h}=x^{0.5}_{i,h} \tag{40}$$

The quantity c is an appropriate positive real-valued constant, and d satisfies (31). When (39) or (40) holds over the larger values of $u_{i,h}$, there does not appear to be any need to limit the maximum value of $u_{i,h}$ to k_2.

VI. Comparison of Systems

Computer-simulation tests have been carried out on the channel estimators described in Sections III–V. The estimators are used in the receiver of a 4800 bit/s QPSK system, operating at 2400 bauds with an 1800 Hz carrier. The various assumptions made in these tests, concerning the data-transmission system and transmission path, are given in Sections I and II. Further details on these are given elsewhere [1], [9], [31].

The results of the tests are given in Tables I-VI. The mean-square error in $Y'_{i,i-1}$ is measured in dB relative to unity, and is given by

$$\xi=10\log_{10}\left(\frac{1}{56000}\sum_{i=4001}^{60000}|Y_i-Y'_{i,i-1}|^2\right) \tag{41}$$

TABLE I
Performance of System 1

ψ	b	θ	ξ
20	0.113	0.966	-18.9
30	0.091	0.949	-25.1
60	0.087	0.934	-31.0

TABLE II
Performance of System 2

Averaging	ψ	b	θ	k_0	k_1	k_2	a	ξ
Growing memory	20	1.0	0.97	0.043	10^{-6}	∞	–	-23.1
	30	1.0	0.96	0.035	10^{-6}	∞	–	-30.2
	60	1.0	0.92	0.030	10^{-6}	∞	–	-41.3
Fading memory	20	1.0	0.97	0.043	10^{-6}	∞	0.25	-23.1
	30	1.0	0.96	0.035	10^{-6}	∞	0.11	-30.2
	60	1.0	0.92	0.030	10^{-6}	∞	0.42	-41.2
Fading memory	20	1.0	0.97	0.043	10^{-6}	∞	1.00	-22.3
	30	1.0	0.96	0.035	10^{-6}	∞	1.00	-29.8
	60	1.0	0.92	0.030	10^{-6}	∞	1.00	-40.9
Growing memory	20	0.9	0.97	0.043	10^{-6}	∞	–	-23.3
	30	1.1	0.96	0.035	10^{-6}	∞	–	-30.2
	60	1.2	0.92	0.030	10^{-6}	∞	–	-41.7
Fading memory	20	1.0	0.97	0.043	10^{-6}	∞	0.25	-23.1
	30	1.1	0.96	0.035	10^{-6}	∞	0.11	-30.2
	60	1.2	0.92	0.030	10^{-6}	∞	0.42	-41.6
Fading memory	20	0.8	0.97	0.043	10^{-6}	∞	1.00	-23.3
	30	1.1	0.96	0.035	10^{-6}	∞	1.00	-29.8
	60	1.2	0.92	0.030	10^{-6}	∞	1.00	-41.0

where the mean-square value of $|Y_i|$ is close to unity. The first 1000 of the received samples in any test are ignored to allow the stabilization of the fading and additive noise processes. During the next 3000 received samples the estimation process operates as described, with a good starting-up procedure, but no measurements are carried out. This eliminates the effect of any transient behavior of the estimator at start up. Over the following 56 000 received samples, ξ is evaluated according to (41). Thus, ξ gives a measure of the steady-state performance of the estimator, which is here taken to be its performance during the prolonged and uninterrupted transmission of the data signal.

The signal/noise ratio is measured as ψ dB where

$$\psi=10\log_{10}(1/(1/2)N_0). \tag{42}$$

Equation (42) uses the fact that the average transmitted energy per bit, at the input and output of the HF radio link, is unity, and the two-sided power spectral density of the additive white Gaussian noise at the output of the HF radio link is $(1/2)N_0$.

In each of Tables I–VI, the adjustable parameters, such as b in (17) or (28), θ in (10) and (11), k_1 and k_2 in (30), d and k_0 in (31), c in (37) or (39), and so on, have been adjusted as far as reasonably possible to minimize ξ. However, in the time available, it has not been possible to carry out the complete

TABLE III
PERFORMANCE OF SYSTEM 3

Averaging	g	ψ	b	θ	d_0	k_1	k_2	a	ξ
Growing memory	31	20	1.0	0.960	4×10^{-5}	10^{-6}	0.120	-	-23.3
		30	1.0	0.947	4×10^{-5}	10^{-6}	0.086	-	-29.0
		60	1.0	0.900	10^{-5}	10^{-6}	0.130	-	-40.6
Fading memory	31	20	1.0	0.960	4×10^{-6}	10^{-6}	0.170	0.04	-23.1
		30	1.0	0.947	3×10^{-6}	10^{-6}	0.086	0.02	-29.0
		60	1.0	0.940	2×10^{-6}	10^{-6}	0.240	0.02	-37.2
Non-Adaptive ($u_{i,h}=1$)	21	20	0.15	0.970	-	-	-	-	-21.1
		30	0.12	0.950	-	-	-	-	-27.8
		60	0.15	0.930	-	-	-	-	-36.8

TABLE IV
PERFORMANCE OF SYSTEM 4

Averaging	ψ	b	θ	c	k_1	k_2	a	ξ
Growing memory	20	1.0	0.985	2.8	10^{-5}	0.72	-	-22.4
	30	1.0	0.977	2.8	10^{-5}	0.64	-	-29.2
	60	1.0	0.954	19.4	0.02	0.48	-	-40.3
Fading memory	20	1.0	0.985	2.8	10^{-5}	0.72	0.95	-21.9
	30	1.0	0.977	2.8	10^{-5}	0.64	0.95	-28.1
	60	1.0	0.954	19.4	0.02	0.48	1.00	-38.0
Fading memory	20	1.0	0.985	2.8	10^{-5}	0.72	1.00	-21.9
	30	1.0	0.977	2.8	10^{-5}	0.64	1.00	-28.1
	60	1.0	0.954	19.4	0.02	0.48	1.00	-38.0

TABLE V
PERFORMANCE OF SYSTEM 5

Averaging	ψ	b	θ	c	k_1	k_2	a	ξ
Growing memory	20	1.0	0.985	0.90	0.001	∞	-	-21.6
	30	1.0	0.976	0.80	0.001	∞	-	-28.7
	60	1.0	0.951	0.72	0.003	∞	-	-40.3
Fading memory	20	1.0	0.985	0.90	0.001	∞	1.0	-22.0
	30	1.0	0.976	0.80	0.001	∞	1.0	-28.8
	60	1.0	0.951	0.72	0.003	∞	1.0	-40.2

TABLE VI
PERFORMANCE OF SYSTEM 6

Averaging	ψ	b	θ	k_0	k_1	k_2	a	ξ
Growing memory	20	1.0	0.985	0.063	10^{-6}	∞	-	-23.6
	30	1.0	0.976	0.064	10^{-6}	∞	-	-30.6
	60	1.0	0.950	0.024	10^{-6}	∞	-	-43.7
Fading memory	20	1.0	0.985	0.063	10^{-6}	∞	0.01	-23.2
	30	1.0	0.976	0.064	10^{-6}	∞	0.01	-30.3
	60	1.0	0.950	0.024	10^{-6}	∞	0.01	-43.4
Fading memory	20	0.86	0.985	0.063	0.001	∞	0.01	-22.3
	30	0.86	0.976	0.064	0.001	∞	0.01	-28.8
	60	0.78	0.950	0.024	0.001	∞	0.01	-40.0
Fading memory	20	1.0	0.985	0.063	10^{-6}	∞	1.0	-23.1
	30	1.0	0.976	0.064	10^{-6}	∞	1.0	-29.5
	60	1.0	0.950	0.024	10^{-6}	∞	1.0	-42.2

optimization of every system, so that it may well be possible to achieve further small improvements in ξ for some systems. In a particular case, for each of Systems 2 and 4 (Tables II and IV), no averaging is applied in the evaluation of $z_{i,h}$ and $x_{i,h}$, respectively, such that $a = 1$ in (26) and (34). Each system is now approximately optimized, subject to $a = 1$ in the fading-memory algorithm. Again, for the first half of the results in Table II, b is fixed at unity.

Three different values of ψ (20, 30, and 60) have been used in the tests, where the values 20 and 30 are such that a significant number of errors in detection of the received data symbols are likely to be caused from time to time by the additive noise, whereas the value 60 represents a high signal/noise ratio, where the fading predominates over the noise.

The good performance achieved by System 2 suggests that the basic mechanism, behind the improvement in performance of Systems 3–6 over System 1, is, at least in part, due to the fact that Systems 3–6 are better able than is System 1 to correct an error in $Y'_{i,i-1}$ caused by an acceleration in Y_i. In System 3, there are a series of local minima in the values of ξ, as the parameter values are varied, instead of a single global minimum. This has led to some difficulty in the selection of the parameter values in Table III.

The rather similar performances of Systems 3–6 suggest that the precise relationship between $u_{i,h}$ and $x_{i,h}$ is not critical, so long as the general form of the relationship does not differ too much from that for System 6. In a practical implementation of any of these systems, $u_{i,h}$ would be determined from $x_{i,h}$ by means of a look-up table, so that the complexity of the relationship between $u_{i,h}$ and $x_{i,h}$ is of no great practical significance.

The growing-memory averages would not be suitable for a practical application of the system, since a drift in phase of the timing waveform at the receiver could introduce considerable changes into the relative peak magnitudes of the different components of Y_i, and, contrary to the case of fading-memory averages, these would not be tracked by the growing-memory averages. The latter have, however, been studied as a check for the effectiveness of the former, because, in the absence of any shift in timing phase or change in fading statistics, the growing-memory averages can be taken to be optimum.

Tests have been carried out with System 6, for two different values of k_1 and also two different values of a (Table VI). A very good performance is obtained here, particularly when $k_1 = 10^{-6}$ and $a = 0.01$. Further tests have been carried out with statistically independent noise components $\{w_i\}$ in (3), in place of the slightly correlated noise components actually obtained at the output of the receiver filters. The tests have been carried out for both Systems 1 and 6, but only a negligibly small difference has been observed. Thus, the correlation in the noise components does not appear to have any significant effect.

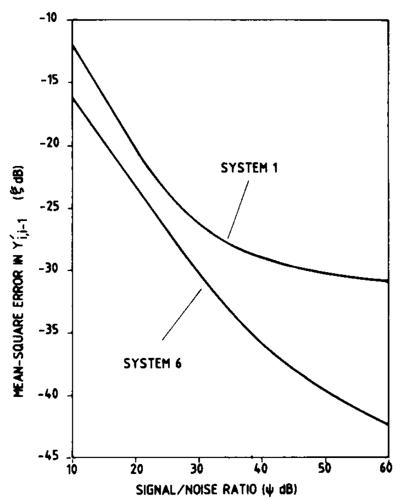


Fig. 8. Comparison of Systems 1 and 6.

In the final set of tests, System 6 (with fading-memory averaging) is compared to System 1, over the whole range of signal/noise ratios $\psi = 10$ to 60, each system being appropriately optimized at each signal/noise ratio. The results of these tests are shown in Fig. 8, which confirms the substantial advantage in performance gained by System 6. The parameter ξ [in (4)] is here evaluated for $i = 6001$ to 60 000, giving a small but not very significant change in performance.

VII. Conclusions

The most promising of the various systems studied here is System 6, which gains advantages over System 1, in tolerance to additive white Gaussian noise, of some 4, 5, and 12 dB, respectively, when $\psi = 20$, 30, and 60. System 6 is not much more complex than System 1, so that it is well worth further study. Clearly, the estimator must be tested with an appropriate near-maximum likelihood detector, to check the effects of errors in detection, and the influence of practical constraints, such as limited arithmetic accuracy, must also be considered.

The fact that System 2 has a performance almost as good as that of System 6 suggests that at least a part of the basic mechanism behind the good performance of System 6 is its ability to track accelerations in Y_i more accurately than can System 1.

References

[1] A. P. Clark and F. McVerry, "Performance of 2400 bit/s serial and parallel modems over an HF channel simulator," in *IERE Conf. Proc. 49*, Loughborough Leics., England, Apr. 1981, pp. 167-179.

[2] J. Pennington, "Comparative measurements of parallel and serial 2.4 kbps modems," *IEE Conf. Publ. No. 206*, pp. 141-144, Feb. 1982.

[3] P. H. Anderson, F. M. Hsu, and M. N. Sandler, "A new adaptive modem for long haul HF digital communications at data rates greater than 1 bps/Hz," in *IEEE Military Commun. Conf.*, Boston, MA, Oct. 1982, vol. 2, pp. 29.2/1-7.

[4] S. Crozier, K. Tiedemann, R. Lyons, and J. Lodge, "An adaptive maximum likelihood sequence estimation technique for wideband HF communications," in *IEEE Military Commun. Conf.*, Boston, MA, Oct. 1982, vol. 3, pp. 29.3/1.9.

[5] A. P. Clark, *Principles of Digital Data Transmission.* London, England: Pentech, 1983, second edition.

[6] W. Hodgkiss and L. F. Turner, "Practical equalization and synchronization strategies for use in serial data transmission over h.f. channels," *Radio Electron. Eng.*, vol. 53, pp. 141-146, Apr. 1983.

[7] W. Hodgkiss, L. F. Turner, and J. Pennington, "Serial data transmission over HF radio links," *IEE Proc.*, part F, vol. 131, pp. 107-116, Apr. 1984.

[8] D. D. Falconer, A. U. H. Sheikh, E. Eleftheriou, and M. Tobis, "Comparison of DFE and MLSE receiver performance on HF channels," *GLOBECOM '83*, San Diego, CA, vol. 1, pp. 13-18, Nov.-Dec. 1983.

[9] A. P. Clark, H. Y. Najdi, and F. McVerry, "Performance of a 9600 bit/s serial modem over a model of an HF radio link," *IEE Conf. Publ. 224*, pp. 151-155, Sept. 1983.

[10] A. P. Clark, *Equalizers for Digital Modems.* London, England: Pentech, 1985.

[11] I. J. Walsh, I. C. Jenkins, E. Hammond, and N. G. Kingsbury, "Implementation of a 2.4 kbits/sec adaptive serial H. F. modem," *IEE Conf. Publ. 245*, pp. 146-150, Feb. 1985.

[12] K. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Trans. ASME J. Basic Eng.*, vol. 83-D, no. 1, pp. 95-108, Mar. 1961.

[13] N. Morrison, *Introduction to Sequential Smoothing and Prediction.* New York: McGraw-Hill, 1969.

[14] A. H. Jazwinski, *Stochastic Process and Filtering Theory.* New York, London: Academic, 1970.

[15] D. Godard, "Channel equalization using a Kalman filter for fast data transmission," *IBM J. Res. & Develop.*, vol. 18, pp. 267-273, May 1974.

[16] P. C. Young, "Recursive approaches to time series analysis," *Bull. Inst. Math. Appl.*, vol. 10, pp. 209-223, May/June 1974.

[17] G. Nicholson and J. P. Norton, "Kalman filter equalization for a time-varying communication channel," *Aust. Telecommun. Res.*, vol. 13, pp. 3-12, 1979.

[18] B. D. O. Anderson and J. B. Moore, *Optimal Filtering.* Englewood Cliffs, NJ: Prentice-Hall, 1979.

[19] F. M. Hsu, "Square root Kalman filtering for high speed data received over fading dispersive HF channels," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 753-763, Sept. 1982.

[20] G. A. Richards, "Implementation of Kalman filters for process identification," *GEC J. of Res.*, vol. 1, no. 2, pp. 100-107, 1983.

[21] J. M. Cioffi and T. Kailath, "Fast recursive least squares transversal filters for adaptive filtering," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, pp. 304-337, Apr. 1984.

[22] J. C. Currie and J. R. Weale, "Studies of inverse filtering and channel estimation techniques for HF serial modems," *IEE Conf. Publ. 245*, pp. 151-155, Feb. 1985.

[23] F. Ling and J. G. Proakis, "Adaptive lattice decision-feedback equalizers—Their performance and application to time-variant multipath channels," *IEEE Trans. Commun.*, vol. COM-33, pp. 348-356, Apr. 1985.

[24] J. M. Cioffi and T. Kailath, "Windowed fast transversal filters adaptive algorithms with normalization," *IEEE Trans. Acoust, Speech, Signal Processing*, vol. ASSP-33, pp. 607-625, June 1985.

[25] A. P. Clark and R. Harun, "Assessment of Kalman-filter channel estimators for an HF radio link," *IEE Proc.*, Part F, vol. 133, pp. 513-520, Oct. 1986.

[26] F. R. Magee and J. G. Proakis, "Adaptive maximum-likelihood sequence estimation for digital signaling in the presence of intersymbol interference," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 120-124, 1973.

[27] A. P. Clark, C. P. Kwong, and F. McVerry, "Estimation of the sampled impulse-response of a channel," *Signal Process.*, vol. 2, pp. 39-53, Jan. 1980.

[28] A. P. Clark and F. McVerry, "Channel estimation for an HF radio link," *IEE Proc.*, Part F, vol. 128, pp. 33-42, Feb. 1981.

[29] ——, "Improved channel estimator for an HF radio link," *Signal Process.*, vol. 5, pp. 241-255, May 1983.

[30] A. P. Clark and S. M. Asghar, "Detection of digital signals transmitted over a known time-varying channel," *IEE Proc.*, Part F, vol. 128, pp. 167-174, June 1981.

[31] A. P. Clark and H. Y. Najdi, "Detection process of a 9600 bit/s serial modem for HF radio links," *IEE Proc.*, Part F, vol. 130, pp. 368-376, Aug. 1983.

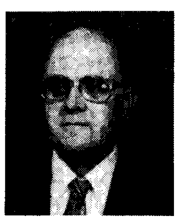
[32] A. P. Clark and S. F. Hau, "Adaptive adjustment of receiver for distorted digital signals," *IEE Proc.*, Part F, vol. 131, pp. 526-536, Aug. 1984.

[33] A. P. Clark and S. N. Abdullah, "Near-maximum likelihood detectors for voiceband channels," *IEE Proc.*, Part F, vol. 134, pp. 217-226, June 1987.

[34] A. P. Clark, Z. C. Zhu, and J. K. Joshi, "Fast start-up channel estimation," *IEE Proc.*, Part F, vol. 131, pp. 375-382, July 1984.

[35] "HF ionospheric channel simulators," *CCIR, 13th Plenary Assembly*, vol. 3, Rep. no. 549, pp. 66-75, 1974.

★



A. P. Clark received the M.A. degree in natural sciences from the University of Cambridge, England, and the Ph.D. degree in electrical engineering from the University of London.

From 1955 to 1965 and again from 1969 to 1970 he worked at Plessey Telecommunications Research (formerly British Telecommunications Research) at Taplow. During this time he was responsible for the basic design, development and testing of a number of different AM, PM, and FM data transmission systems for use over the telephone network. He was also involved in the design of error detecting codes, the investigation of modulation and multiplex methods for digital signals and the study of adaptive detection processes. From 1965 to 1968 he was in the Department of Electrical Engineering of the Imperial College of Science and Technology, and was engaged in research on random access discrete address communication systems. Since 1970 he has been in the Department of Electronic and Electrical Engineering of Loughborough University of Technology where he is now Professor of Telecommunications. His current research interests include the study of techniques for maximizing the transmission rate of digital systems over band-limited channels, with particular emphasis on satellite and radio links.

★

S. Hariharan received the B.S. degree in electrical engineering in 1976 and M.Tech. in microwave electronics in 1978 from the University of Delhi, India.

From 1978 to 1985 he worked with the Telecommunication Research Centre, New Delhi, India, as a Research and Development Engineer in the area of Digital Microwave communication. He is presently a Commonwealth Scholar of research at the Department of Electronic and Electrical Engineering, University of Technology, Loughborough, UK. His research interests include detection and estimation theory, digital signal processing and digital microwave communication.