Square Root Kalman Filtering for High-Speed Data Received over Fading Dispersive HF Channels

FRANK M. HSU, MEMBER, IEEE

Abstract—This paper shows how to apply square root Kalman filtering to high speed data transmission on fading dispersive high-frequency (HF) channels. High data-rate HF transmission can be achieved by use of a single carrier bandwidth efficiency modulation that is demodulated with decision-feedback equalization (DFE). The DFE coefficients are updated adaptively by square root Kalman algorithms. Theoretical formulation and mechanization procedures for the square root Kalman algorithms are given. Computer simulation indicates that good error-rate performance can be achieved with these algorithms for rapid fading HF radio channels.

I. Introduction

THIS PAPER is concerned with square root Kalman algorithms for adaptive equalization of fading dispersive high-frequency (HF) radio channels. High-speed adaptive communication systems are constrained by the nonideal characteristics of HF channels such as bandwidth restrictions, signal fading, and multipath dispersion. These channel limitations present the major obstacles to achieving increased digital transmission rates with the required accuracies.

High data-rate HF transmission (up to 9600 bits/s in 3 KHz bandwidths) can be achieved by use of a single carrier bandwidth-efficient modulation that is demodulated with a decision feedback equalizer (DFE) operated adaptively [1]. The DFE coefficients can be updated adaptively by gradient, Kalman algorithms, etc. It is found [2] that only the Kalman algorithm provides a tracking rate sufficient to follow the time-varying HF channel. Unfortunately, the Kalman algorithm has been found to be sensitive to computer roundoff errors; numerical accuracy due to roundoff sometimes degrades performance to the point where the results become meaningless [3]. Simulations that we have conducted in fixed-point arithmetic have shown the performance of a DFE with conventional Kalman updating does indeed degrade rapidly as the computer-word size is decreased. Recently, Falconer and Ljung [4] have shown a "fast Kalman algorithm" which requires only n rather than n^2 operations per cycle. In fact, the fast Kalman algorithm was found, independently, by us and by Lim and Mueller [5] to be unstable (worse than conventional Kalman algorithm) in time-varying channel environments. In other applications, researchers have shown that square root formulations of the Kalman algorithm have inherently better stability and numerical accuracy than does the conventional Kalman algorithm [3], [6]–[13]. The improved numerical behavior of the square root algorithms is due in large part to a reduction of the numerical ranges of the variables. Loosely speaking, one can say that computations which involve numbers ranging between 10^{-N} and 10^{N} are reduced to ranges between $10^{-N/2}$ and $10^{N/2}$. Thus the square root algorithm achieve accuracies that are comparable with a conventional Kalman algorithm that uses twice the numerical precision [3].

The application of conventional Kalman filtering theory to the adaptive equalization techniques has been considered by Godard [14]. In this paper, new algorithms based on the Kalman/Godard algorithm and the Carlson/Bierman [3], [6], [12] (or Gentleman [13]) *U-D* covariance factorization filter are proposed for updating the tap gains of a decision-feedback equalizer for fading dispersive HF channels. These new algorithms are applicable to complex (i.e., angle-modulated) input signals and time-varying channels. Theoretical formulation and mechanization procedures for the square root Kalman algorithms are given here. Computer simulation indicates that good error-rate performance can be achieved by these algorithms for rapidly fading HF radio channels.

The computational requirements of the square root algorithms for an equalizer with N taps are 1.5 N^2 complex multiplications per cycle, which are roughly equivalent to the requirements of the conventional Kalman algorithm. Typically, the keying rate of an HF modem with 3 KHz transmission bandwidth is set to 2400 bits/s. A decision feedback equalizer with 15 feedforward taps and 14 feedback taps can handle multipath time spread up to 6.25 ms. The multipath spread is less than 5 ms in typical HF channels [15]. The least-squares lattice structure introduced by Morf *et al.* and adapted by Satorius and Park [16] for equalizer adjustment algorithm is a linear (all zero) equalization technique. It is well known [17] that linear filters are unable to cope with severe fading-dispersive channels. The

Manuscript received March 25, 1980; revised July 6, 1981.
The author is with the Sylvania Systems Group, Communications Systems Division, GTE Products Corporation, 77 A Street, Needham Heights, MA 02194.

0018-9448/82/0900-0753$00.75

least squares lattice decision-feedback equalizer formulated by Shensa [18] requires about 97 N multiplications per cycle, where N is the number of stage (in a N-zero, N-pole channel). The least-squares lattice DFE requires more computation than the square root algorithms shown in this paper for multipath time spreads up to 6.25 ms. Consequently, DFE's with square root Kalman algorithms are more efficient than DFE's with least-squares lattice algorithm for HF modem implementation. The square root Kalman algorithms described are shown to be remarkably stable with rapid convergence and tracking abilities for short computer-word sizes. The implementation of these algorithms for high-speed HF modems has been accomplished recently [19].

II. Communications System Model

The HF channel can be modeled as a tapped delay-line with time-varying coefficients followed by addition of zero-mean white-Gaussian noise, $\nu(t)$ with variance, σ_n^2 [13]. The tapped delay-line has taps spaced at no more than the symbol interval T and nonzero tap coefficients corresponding to discrete multipath propagation. The discrete tapped delay-line model can combine the transmission filter, the channel, the matched filter, the symbol rate samples and a noise whitening filter [1]. The tap coefficients are assumed to be independent complex-Gaussian random variables $\{f_i,\ i = 0,\cdots,L\}$ with zero means and fixed variances, $E\{|f_i|^2\}$. The tap coefficients are generated by filtering white-Gaussian noise through a filter whose bandwidth is on the order of the fade rate.

It can be shown that passage of the random signals through the discrete tapped delay-line results in a received signal sequence $\{R_k\}$ which can be expressed as

$$R_k = \sum_{m=0}^{L} f_m I_{k-m} + \nu_k, \qquad k = 0,1,2,\cdots, \tag{2.1}$$

where $\{I_k\}$ is the random signal sequence. In order to reduce the signal fluctuations due to the fading channel, the received signal is normally regulated by an automatic gain control (AGC) filter that keeps the magnitude of the output signal y_k nearly constant:

$$y_k = \frac{R_k}{\sqrt{H_k}}, \tag{2.2}$$

$$H_k = \lambda R_k R_k^* + (1-\lambda) H_{k-1}, \tag{2.3}$$

where λ is the AGC constant and asterisk (*) denotes complex conjugate. The signal y_k can be demodulated with a decision-feedback equalizer (DFE) operated adaptively. The DFE consists of two sections, a feedforward section and a feedback section. The feedforward section is a linear transversal filter with tap coefficients $\{a_j,\ j = 0,\cdots,M\}$ spaced at no more than the symbol interval T. The input to the feedforward section is the sequence $\{y_k\}$ of noisy received signal samples. The feedback section is also a linear transversal filter with tap coefficients $\{b_j,\ j = 1,\cdots,N-M-1\}$ spaced at the symbol interval. The input to the feedback section is the output symbol decision sequence from the nonlinear symbol detector. The equalizer output can be expressed as

$$\hat{I}_k = \sum_{j=0}^{M} a_j y_{k+j} + \sum_{j=1}^{N-M-1} b_j \tilde{I}_{k-j}, \tag{2.4}$$

where $\hat{I}_k$ is an estimate of the kth information symbol and $\{\tilde{I}_{k-1},\cdots,\tilde{I}_{k-N+M+1}\}$ are previously detected symbols. The decision $\tilde{I}_k$ is formed by quantizing the estimate $\hat{I}_k$ to the nearest information symbol. The least mean-square error (LMSE) criterion provides a practical means for selecting the equalizer tap coefficients $\{a_j\}$ and $\{b_j\}$. Based on the assumption that previously detected symbols in the feedback section are correct, the minimization of the mean square error

$$\mathcal{E}^2 = E\,|I_k - \hat{I}_k|^2 \tag{2.5}$$

leads to the following set of linear equations for the tap coefficients of the feedforward filter

$$\sigma_s^2 f_m'^* = \sum_{j=0}^{M} \sum_{l=0}^{m} a_j \left[\sigma_s^2 f_l'^* f_{j-m+l}' + \sigma_{in}^2 \delta_{jm}\right], \qquad m = 0,1,\cdots,M, \tag{2.6}$$

$$f_k' = \frac{f_k}{\sqrt{H_k}}, \tag{2.7}$$

$$n_k = \frac{\nu_k}{\sqrt{H_k}}, \tag{2.8}$$

where σ_s^2 and σ_{in}^2 are the variances of the signal I_k and noise n_k, respectively. The tap coefficients of the feedback filter are given in terms of coefficients for the feedforward filters by the following expression

$$b_m = -\sum_{j=0}^{M} a_j f'_{m+j}, \qquad m = 1,2,\cdots,N-M-1. \tag{2.9}$$

One can easily show that these values of the feedback coefficients result in complete elimination of intersymbol interference from previously detected symbols, provided that previous decisions are correct. The optimum number $N-M-1$ of feedback taps may be set equal to the number L of interfering symbols given in the discrete channel model, if the tap a_0 of the feedforward filter is selected as the main tap for synchronization purposes.

The equalizer tap coefficients $\{a_j\}$ and $\{b_j\}$ may be computed exactly from (2.6) and (2.9) provided the values of σ_s^2, σ_{in}^2, and $\{f_l'\}$ are known. Normally, however, tap coefficients are computed adaptively because the channel is not known *a priori* at the receiver. The calculation of the theoretical tap coefficients from (2.6) and (2.9) is primarily useful in providing a baseline performance level for comparison with the various adaptive schemes.

III. Kalman/Godard Filter

The solution of (2.6) and (2.9) for the optimum tap coefficients involves the inversion of a matrix that depends on the signal and noise statistics as well as the characteristics of the channel. Generally, this is not known to the receiver so that iterative procedures must be used to determine the tap coefficients. The Kalman/Godard filter is designed to update the equalizer tap coefficients, adaptively, from the received signal samples. During the training period, the equalizer continuously seeks to minimize the deviation of its sampled output signal from an ideal reference signal internally generated by the receiver. When the residual distortion is small enough, the equalizer is switched into the decision-directed mode, using as a reference a reconstructed signal obtained by quantizing the output signal of the equalizer to the nearest symbol. Then the equalizer has the ability to self-adapt to changes in channel characteristics during the transmission process. The Kalman/Godard algorithm will be described here for updating the DFE coefficients. This algorithm is by far the most powerful, i.e., it is fastest, known adaptive algorithm for adjusting the equalizer coefficients.

At time k, the inputs $\{y_k, y_{k+1}, \cdots, y_{K+M}\}$ and decisions $\{\tilde{I}_{k-1}, \tilde{I}_{k-2}, \cdots, \tilde{I}_{k-N+M+1}\}$ are in the feedforward and feedback sections of decision feedback equalizer, respectively. The signals can be represented by a vector

$$X(k)^t = \big(y_k, y_{k+1}, \cdots, y_{k+M}, \cdot \tilde{I}_{k-1}, \tilde{I}_{k-2}, \cdots, \tilde{I}_{k-N+M+1}\big), \quad (3.1)$$

where t denotes the transpose.

The equalizer tap coefficients at time k are represented by the vector

$$C(k-1)^t = (a_0, a_1, \cdots, a_M, b_1, b_2, \cdots, b_{N-M-1}), \quad (3.2)$$

and its output is $C(k-1)^t X(k)$, which may differ from the ideal output I_k by an error $\epsilon(k)$

$$\epsilon(k) = I_k - C(k-1)^t X(k). \quad (3.3)$$

The goal of the adaptation algorithm is to adjust $C(k)$ iteratively to converge toward an optimum vector $C(k)_{\text{opt}}$, which minimizes the mean-squared value of $\epsilon(k)$. The new estimate of $C(k)$ is given by

$$C(k) = C(k-1) + G(k)\epsilon(k), \quad (3.4)$$

where

$$G(k) = \big[X(k)^t P'(k) X^*(k) + \xi'\big]^{-1} P'(k) X^*(k), \quad (3.5)$$

$$P(k) = P'(k) - G(k) X(k)^t P'(k), \quad (3.6)$$

$$P'(k) = \Phi(k, k-1) P(k-1) \Phi^{*t}(k, k-1) + Q(k). \quad (3.7)$$

All constants, vectors and matrices (except ξ') defined in (3.1), (3.2), (3.3), (3.4), (3.5), (3.6), and (3.7) are complex quantities. The quantities a_k, b_k, I_k, and $\epsilon(k)$ are scalars; $X(k)$, $C(k)$, and $G(k)$ are vectors; and $P(k)$, $P'(k)$, $\Phi(k, k-1)$, and $Q(k)$ are square matrices. The asterisk and superscript t denote complex conjugate and transpose respectively. The matrix $P(k)$ represents the covariance of the error in the estimate of the actual equalizer coefficients.

The state transition matrix $\Phi(k, k-1)$ is defined in a system modeled as

$$C(k) = \Phi(k, k-1) C(k-1) + W(k), \quad (3.8)$$

where $W(k)$ is N-dimensional vector of zero-mean white-noise process. It is assumed that the noise processes W and ϵ are statistically independent. The parameter ξ' denotes the approximate final mean square error, $E[|\epsilon|^2]$. The matrix $Q(k)$ represents the covariance matrix of W. Generally, the state transition matrix is not known to the receiver in the equalization process. Nevertheless, one can assume that the optimum tap-gain values are randomly varying about a mean value [14], i.e.,

$$C(k)_{\text{opt}} = C(k-1)_{\text{opt}} + \Delta C(k), \quad (3.9)$$

where $\Delta C(k)$ is considered to be a noise process. The matrix $Q(k)$ is then reduced to

$$Q(k) = E\big[(\Delta C(k))(\Delta C(k))^{*t}\big] \quad (3.10)$$

and, at each step from (3.7), the predicted-error covariance matrix is given by

$$P'(k) = P(k-1) + Q(k). \quad (3.11)$$

Exact calculation of the correlation matrix $Q(k)$ is impossible because the optimum tap coefficients are unknown to the receiver. It has been shown [14] that $C(k)$ converges to $C(k)_{\text{opt}}$ within less than $2N$ steps. If the matrix $Q(k)$ is assumed to be negligible for a slowly time-varying channel, then one has [14]

$$P(k-1) = \xi' \left[\sum_{m=0}^{k-1} X^*(m) X(m)^t\right]^{-1} \quad (3.12)$$

and

$$E\big[X^*(k) X(k)^t\big] = \xi' k^{-1} P(k-1)^{-1}, \quad (3.13)$$

as k goes to infinity. It can be verified [14] that the $P(k)$ (or $P'(k)$) matrix elements converge to zero. Consequently, the parameter $\alpha' = X(k)^t P'(k) X^*(k) + \xi'$ converges to ξ'. In the steady-state limit, the Kalman gains in (3.5) are reduced to

$$G(k) = P(k-1) X^*(k)/\xi'. \quad (3.14)$$

Substituting (3.4), (3.13), and (3.14) into (3.10) yields

$$Q(k) = k^{-1} P(k-1). \quad (3.15)$$

The matrix $Q(k)$ is identically zero in the case of stationary channels. For slowly time-varying channels, $Q(k)$ may be assumed proportional to $P(k-1)$, $Q(k) = qP(k-1)$, where q ($\ll 1$) is a small constant. Substituting (3.15) into (3.11), (3.5), and (3.6), we obtain

$$G(k) = \big[X(k)^t P(k-1) X^*(k) + \xi\big]^{-1} P(k-1) X^*(k), \quad (3.16)$$

$$P(k) = (1+q)\big[P(k-1) - G(k) X(k)^t P(k-1)\big], \quad (3.17)$$

where

$$\xi = \frac{\xi'}{1+q}. \tag{3.18}$$

The parameter ξ and q should be adjusted to obtain optimum performance. It is interesting to note that (3.16) and (3.17) have the form that would result from deriving the Kalman updating algorithm to minimize the exponentially weighted square error at time k:

$$\mathcal{E}^2(k) = \sum_{n=1}^{k} w^{k-n} |I_n - C(k)^t X(n)|^2, \tag{3.19}$$

Specifically, (3.16) and (3.17) are the solution to this minimization problem if we let $(1+q)^{-1} = w$ and $\xi = w$. The exponentially weighted error criterion has intuitive appeal in that the distant past is "forgotten", which is a desirable property for following time-variations of the channel and helping avoid problems associated with digital-roundoff errors. This method has been investigated by Satorious [16] and Shensa [18], etc. Unfortunately, the optimum values of q and ξ are quite different from that specified in the exponentially weighted method.

The starting conditions for $C(k)$ and $P(k)$ are given by

$$C(0) = 0, \tag{3.20}$$

$$P(0) = I, \tag{3.21}$$

where 0 is the zero vector and I is the identity matrix.

IV. Square Root Kalman Filter

The classic Kalman filter presented in Section III provides a tracking rate sufficient for fading HF radio channels, thereby achieving good error-rate digital communication performance. Unfortunately the covariance update formula (3.6) is numerically unstable. The main reason for this instability is that P is computed as a difference of two positive semidefinite matrices. Numerical accuracy is reduced in every iteration. Moreover, numerical deterioration is also associated with high dimensionality of G and P on X and with the accumulated effects of roundoff error. The accuracy degeneration of the classic Kalman filter may result in a P matrix which is indefinite (having both positive and negative eigenvalues).

The Carlson/Bierman U-D covariance factorization filter [3], [6], [12] involves an upper triangular factorization of a real-valued filter error covariance matrix, i.e., $P = UDU^t$. This algorithm can be linked to square root Kalman filtering, although the computation of square roots is not required. The U-D factorization procedure guarantees nonnegativity of the computed covariance matrix. However, the Carlson/Bierman algorithm cannot be directly applied to HF communication systems. First, the baseband signals in the equalizer are generally represented by complex variables. Second, the correlation matrix Q shown in (3.7), is not included in the Carlson/Bierman square root Kalman algorithm. Godard suggested the need for this modification in order to apply the Kalman algorithm to time-varying channels [14].

The algorithm presented in this section gets around both of the above limitations. It involves an upper complex triangular factorization of the filter error covariance matrix, i.e., $P = U^*DU^t$. Efficient and stable measurement-updating recursions are achieved for the upper triangular matrix U and the diagonal matrix D. This method is based on (3.17) in which the correlation matrix Q is computed implicitly in each iteration. The details of the algorithm will be given in the following discussion.

The price to be paid in direct implementation of (3.17) is quite high; however, the square root factorization algorithm allows the multiplications by $(1+q)$ to be incorporated into a simple (real-valued) diagonal matrix. This is accomplished as follows.

Suppose that the error covariance matrix P is expressed in factored form as

$$P(k) = U^*(k)D(k)U^t(k), \tag{4.1}$$

where U is an upper triangular matrix with unit diagonal elements and nonequal off-diagonal elements μ_{ij}, $i = 1, 2, \cdots, N-1$; $j = i+1, i+2, \cdots, N$, and D is a diagonal matrix with real-valued diagonal elements $(d_1, d_2, \cdots, d_N)$. Also let the vectors F and V be defined by

$$F(k-1) = U^t(k-1)X^*(k) \tag{4.2}$$

and

$$V(k-1) = D(k-1)F(k-1). \tag{4.3}$$

Substituting (3.16), (4.1), (4.2), and (4.3) into (3.17) yields

$$U^*(k)D(k)U^t(k) = (1+q)U^*(k-1)\{D(k-1) - \alpha^{-1}V(k-1)V^{*t}(k-1)\}U^t(k-1), \tag{4.4}$$

where $\alpha = X(k)^t P(k-1)X^*(k) + \xi$. If we define $\bar{U}$ and $\bar{D}$ factors of the terms inside the braces of (4.4), we obtain

$$\bar{U}(k-1)^*\bar{D}(k-1)\bar{U}^t(k-1) = D(k-1) - \alpha^{-1}V(k-1)V^{*t}(k-1), \tag{4.5}$$

and

$$U^*(k)D(k)U^t(k) = (1+q)(U(k-1)\bar{U}(k-1))^* \cdot \bar{D}(k-1)(U(k-1)\bar{U}(k-1))^t. \tag{4.6}$$

From (4.6) we can identify

$$U(k) = U(k-1)\bar{U}(k-1) \tag{4.7}$$

and

$$D(k) = (1+q)\bar{D}(k-1). \tag{4.8}$$

Thus the updated U-D factors are determined in terms of the U-D factors of $D(k-1) - \alpha^{-1}V(k-1)V^{*t}(k-1)$ and $U(k-1)$.

Theorem 1: Let the factors X, F, and V be denoted as

$$X^t(k) = (x_1, x_2, \cdots, x_N), \tag{4.9}$$

$$F^t(k-1) = (f_1, f_2, \cdots, f_N), \tag{4.10}$$

and

$$V^t(k-1) = (v_1, v_2, \cdots, v_N), \tag{4.11}$$

then the U-D factors of $D(k-1) - \alpha^{-1}V(k-1)V^{*t}(k-1)$ can be obtained by evaluating the following ordered equations recursively for $j = 2,\cdots,N$:

$$\bar{d}_1 = d_1(k-1)\frac{\xi}{\alpha_1}, \tag{4.12}$$

$$\bar{d}_j = d_j(k-1)\frac{\alpha_{j-1}}{\alpha_j}, \tag{4.13}$$

and for $i = 1,2,\cdots,j-1$,

$$\bar{\mu}_{ij} = -\frac{v_i{}^* v_j}{\alpha_j \bar{d}_j}, \tag{4.14}$$

where

$$f_1 = x_1{}^*, \tag{4.15}$$

$$f_j = \sum_{i=1}^{j-1} \mu_{i,j}(k-1)x_i{}^* + x_j{}^*, \qquad j = 2,3,\cdots,N, \tag{4.16}$$

$$v_j = d_j(k-1)f_j, \qquad j = 1,2,\cdots,N, \tag{4.17}$$

$$\alpha_1 = \xi + v_1 f_1{}^*, \tag{4.18}$$

$$\alpha_j = \alpha_{j-1} + v_j f_j{}^*, \qquad j = 2,3,\cdots,N. \tag{4.19}$$

Theorem 2: The intermediate Kalman gain G and the updated error covariance factors U and D can be obtained from the following algorithm

$$g_j = v_j, \qquad j = 1,2,\cdots,N, \tag{4.20}$$

$$d_1(k) = (1+q)\bar{d}_1, \tag{4.21}$$

$$d_j(k) = (1+q)\bar{d}_j, \qquad j = 2,3,\cdots,N, \tag{4.22}$$

$$\lambda_j = -f_j/\alpha_{j-1}, \qquad j = 2,3,\cdots,N, \tag{4.23}$$

$$\mu_{ij}(k) = \mu_{ij}(k-1) + g_i{}^*\lambda_j, \qquad i = 1,2,\cdots,j-1, \tag{4.24}$$

$$g_i = g_i + v_j \mu_{ij}^*(k-1), \qquad i = 1,2,\cdots,j-1, \tag{4.25}$$

where g_i on the right-hand side of (4.25) is g_i obtained from $(j-1)$th iteration, and

$$G^t = (g_1, g_2, \cdots, g_N), \tag{4.26}$$

and the final Kalman gain is given by

$$G(k) = G/\alpha_N. \tag{4.27}$$

The square root Kalman algorithm is given by Theorems 1 and 2.

V. A Revised Square Root Formulation

The square root algorithm described in the previous section is a stable algorithm as long as the parameters ξ and q are well chosen. ξ and q are sensitive to the rate of convergence of equalization and fading rate of the channel, respectively. Intuitively, the value of q after equalization should be proportional to the magnitude of the fading rate. The optimum value of q may also depend on the algorithm utilized for updating the tap gains. On the other hand, the convergence rate of the algorithm is primarily controlled by the parameter ξ. A large value of ξ will produce a slow rate of stable convergence, whereas a small value of ξ will result in a fast rate of convergence but at the sacrifice of stability. From (4.9) to (4.22), it is found that $d_1(1)$ is significantly different from $d_j(1)$, $j = 2,3,\cdots,N$ for small values of ξ(e.g., $\xi = 0.001$; $\alpha_1 \approx 1$; $d_1(1) \approx 0.001$; $d_j(1) \approx 1$, $j = 2,3,\cdots,N$; $q \ll 1$). The magnitude disparity of d_1 and d_j in the start-up period can significantly disturb the equalization process. The disparity can also be expected to degrade the performance of the algorithm if a periodic resetting scheme as described in Section VII is adopted. The algorithm can be modified in such a way that it is stable for various values of ξ without sacrificing the rate of convergence. A revised square-root formulation described below will serve this purpose. This algorithm is more suited for implementing the receiver with a processor of limited word size.

Theorem 3: If (3.6) is modified according to the following equation

$$P(k) = P(k-1) + Q(k) - G(k)X(k)^t P(k-1), \tag{5.1}$$

then the intermediate Kalman gain G and the updated error covariance factors U and D can be obtained from the following algorithm

$$g_j = v_j, \qquad j = 1,2,\cdots,N, \tag{5.2}$$

$$d_1(k) = d_1(k-1)h_q\frac{(\xi + h_t)}{(\alpha_1 + h_t)}, \tag{5.3}$$

$$d_j(k) = d_j(k-1)h_q\frac{(\alpha_{j-1} + h_t)}{(\alpha_j + h_t)}, \qquad j = 2,3,\cdots,N, \tag{5.4}$$

$$\lambda_j = -\frac{f_j}{(\alpha_{j-1} + h_t)}, \qquad j = 2,3,\cdots,N, \tag{5.5}$$

$$\mu_{ij}(k) = \mu_{ij}(k-1) + g_i{}^*\lambda_j, \qquad i = 1,2,\cdots,j-1, \tag{5.6}$$

$$g_i = g_i + v_j\mu_{ij}^*(k-1), \qquad i = 1,2,\cdots,j-1, \tag{5.7}$$

where g_i on the right-hand side of (5.7) is g_i obtained from $(j-1)$th iteration, and

$$f_1 = x_1{}^*,$$

$$f_j = \sum_{i=1}^{j-1} \mu_{ij}(k-1)x_i{}^* + x_j{}^*, \qquad j = 2,3,\cdots,N, \tag{5.9}$$

$$v_j = d_j(k-1)f_j, \qquad j = 1,2,\cdots,N, \tag{5.10}$$

$$\alpha_1 = \xi + v_1 f_1{}^*, \tag{5.11}$$

$$\alpha_j = \alpha_{j-1} + v_j f_j{}^*, \qquad j = 2,3,\cdots,N, \tag{5.12}$$

$$h_t = \alpha_N q, \tag{5.13}$$

$$h_q = 1 + q, \tag{5.14}$$

$$G^t = (g_1, g_2, \cdots, g_N), \tag{5.15}$$

and the final Kalman gain is given by

$$G(k) = G/\alpha_N. \tag{5.16}$$

Proof: Let F and V be given in (4.2) and (4.3), the U-D factors of P given in (5.1) become

$$U^*(k)D(k)U^t(k) = U^*(k-1)\{D'(k-1) - \alpha^{-1}V(k-1)V^{*t}(k-1)\}U^t(k-1), \tag{5.17}$$

where D' is a diagonal matrix with diagonal elements $((1+q)d_1, (1+q)d_2, \cdots, (1+q)d_N)$. With extensive mathematical manipulations (see Appendix), it can be shown that the U-D factors of terms inside the braces of (5.17)

$$\bar{P} = D'(k-1) - \alpha^{-1}V(k-1)V^{*t}(k-1) \tag{5.18}$$

are given by

$$\bar{d}_1 = d_1(k-1)h_q\frac{(\xi + h_t)}{(\alpha_1 + h_t)}, \tag{5.19}$$

$$\bar{d}_j = d_j(k-1)h_q\frac{(\alpha_{j-1} + h_t)}{(\alpha_j + h_t)}, \qquad j = 2,3,\cdots,N, \tag{5.20}$$

$$\bar{\mu}_{ij} = -\frac{h_q v_i^* v_j}{(\alpha_j + h_t)\bar{d}_j}, \qquad i = 1,2,\cdots,j-1, \tag{5.21}$$

where the variables f_j, v_j, α_j, h_t, and h_q are given in (5.8) to (5.14). Then the U-D factors of P follow immediately from Theorem 2, i.e.,

$$d_1(k) = \bar{d}_1, \tag{5.22}$$

$$d_j(k) = \bar{d}_j, \qquad j = 2,3,\cdots,N, \tag{5.23}$$

$$\lambda_j = -\frac{f_j}{(\alpha_{j-1} + h_t)}, \qquad j = 2,3,\cdots,N, \tag{5.24}$$

$$\mu_{ij}(k) = \mu_{ij}(k-1) + g_i^*\lambda_j, \qquad i = 1,2,\cdots,j-1, \tag{5.25}$$

$$g_i = g_i + v_j\mu_{ij}^*(k-1), \qquad i = 1,2,\cdots,j-1, \tag{5.26}$$

where

$$G^t = (g_1, g_2, \cdots, g_N), \tag{5.27}$$

and the final Kalman gain is given by

$$G(k) = G/\alpha_N. \tag{5.28}$$

Q.E.D.

The revised square root formulation is stable in terms of various values of ξ. A small value of ξ could be used for a brief portion of the start-up phase to achieve a fast reduction of mean-square distortion. However, for a fading channel, four computer runs with $\xi = 1.0$, 0.1, 0.01, and 0.001 gave identical results for the same sequence $\{I_k\}$ and the same sequence of noise samples n_k after equalization.

VI. Computational Procedures

Significant computer storage can be saved for numerical computation of the square root Kalman algorithms. The computational procedures described here are such that the lower triangular portion and the diagonal elements of U are not used (the diagonal elements of U are unity). The vector V is replaced by vector G for storage reduction. The computational procedures of the square root (U-D factorization) formulation defined in Section IV are given by

$$\epsilon(k) = \tilde{I}_k - \sum_{j=1}^{N} x_j c_j, \tag{6.1}$$

$$f_1 = x_1^*, \tag{6.2}$$

$$f_j = \sum_{i=l}^{j-1} \mu_{i,j}x_i^* + x_j^*, \qquad j = 2,3,\cdots,N, \tag{6.3}$$

$$g_j = d_j(k-1)f_j, \qquad j = 1,2,\cdots,N, \tag{6.4}$$

$$\alpha_1 = \xi + g_1 f_1^*, \tag{6.5}$$

$$\alpha_j = \alpha_{j-1} + g_j f_j^*, \qquad j = 2,3,\cdots,N, \tag{6.6}$$

$$h_q = 1 + q, \tag{6.7}$$

$$\gamma = \frac{1}{\alpha_1} \tag{6.8}$$

$$d_1(k) = d_1(k-1)h_q\xi\gamma, \tag{6.9}$$

$$\beta = \alpha_{j-1}, \tag{6.10}$$

$$\lambda_j = -f_j\gamma, \tag{6.11}$$

$$\gamma = \frac{1}{\alpha_j}, \tag{6.12}$$

$$d_j(k) = d_j(k-1)h_q\beta\gamma, \tag{6.13}$$

$$\beta_1 = \mu_{ij}, \tag{6.14}$$

$$\mu_{ij} = \beta_1 + g_i^*\lambda_j, \tag{6.15}$$

$$g_i = g_i + g_j\beta_1^*. \tag{6.16}$$

All constants, except ξ, α_j, $j = 1,2,\cdots,N$, γ, h_q, q, d_j, $j = 1,,\cdots,N$, and β defined in (6.1) to (6.16), are complex quantities. The quantities defined in (6.1) to (6.9) are computed first. Equations (6.10) to (6.16) are evaluated recursively for $j = 2,3,\cdots,N$, and (6.14) to (6.16) are computed in order for $i = 1,2,\cdots,j-1$. The equalizer tap weights $C^t = (C_1, C_2, \cdots, C_N)$ are given by

$$e = \epsilon(k)\gamma, \tag{6.17}$$

$$C_j = C_j + g_j e, \qquad j = 1,2,\cdots,N, \tag{6.18}$$

where the computation of the final Kalman gain, $g_j = g_j\gamma$, $j = 1,2,\cdots,N$, is avoided by adding an intermediate step defined in (6.17). Equations (6.1) to (6.18) are evaluated for every new input signal X defined in (3.1).

For the revised square root formulation, the computational procedures are given in (6.1) to (6.18) except (6.8),

TABLE I
COMPUTATIONAL REQUIREMENTS FOR SQUARE ROOT KALMAN ALGORITHMS

Algorithms	Equivalent Real Multiplications	Equivalent Real Additions	Real Reciprocals	Computer Storage
Square Root Algorithm	$6N^2 + 11N$	$6N^2 + 6N$	N	$N^2 + 8N + 14$
Revised Square Root Algorithm	$6N^2 + 11N + 1$	$6N^2 + 7N + 1$	$N + 1$	$N^2 + 8N + 16$

(6.9), (6.10), and (6.12) are modified as

$$\gamma = \frac{1}{\alpha_1 + h_t}, \tag{6.19}$$

$$d_1(k) = d_1(k-1)h_q(\xi + h_t)\gamma, \tag{6.20}$$

$$\beta = \alpha_{j-1} + h_t, \tag{6.21}$$

$$\gamma = \frac{1}{\alpha_j + h_t}, \tag{6.22}$$

where $h_t = \alpha_N q$.

The computational requirements of the square root algorithms for an equalizer with N taps are listed in Table I. The requirements may be slightly different for various signal processors.

VII. SIMULATION RESULTS

In this section we present some results of computer simulation of the properties and performance of the algorithms presented in the previous sections. We have used the fading dispersive HF radio channel model described previously. A complex Gaussian noise is also generated and added to the received signal.

This section contains the properties of convergence, theoretical and simulated performance results for the DFE with 8-ary $(8-\phi)$ phase-shift keying (PSK) modulation. Extensive simulations were conducted in the investigation of the performance of the receiver using the DFE with square root Kalman algorithms for mitigating the intersymbol interference and signal fading. The channel fade rate is 1 Hz and the channel model consists of two fading paths (except in Fig. 3). The received signal is regulated by an AGC filter in the front end with an AGC constant, $\lambda = 0.02$. The DFE consists of two sections, a feedforward section and a feedback section. The numbers of feedforward and feedback taps are set to be greater and equal to the number of interfering symbols given in the discrete channel model, respectively. The initial values of U and D in the square root algorithms are

$$d_j = 1.0, \quad j = 1, 2, \cdots, N, \tag{7.1}$$

$$\mu_{ij} = (0.0, 0.0), \quad i = 1, 2, \cdots, N; \quad j = i+1, i+2, \cdots, N, \tag{7.2}$$

where x, y, and (x, y) denote the real part, imaginary part and complex quantities.

Figs. 1 and 2 depict the convergence properties of the equalizer tap weight with a two-path fading channel model in the first 40 iterations. The DFE with square root and revised square root Kalman algorithms converges very rapidly following the initial start-up. The speed of convergence and stability of the algorithms will depend on the parameters λ, ξ, and q chosen for various design objectives.

Fig. 3 shows the convergence properties of the equalizer tap-weight with a two-path fixed channel using the revised square root Kalman algorithm. It is clear that the convergence of a DFE with the revised square root algorithm is extremely fast and stable on fixed channels. These results indicate that the square root Kalman algorithms can also be used to obtain high stability and fast start-up on telephone channels.

Figs. 4 and 5 depict the performance of the DFE for $8\ \phi$ PSK with a two-path fading channel model using the square root and revised square root Kalman algorithms. The quantized symbol $\tilde{I}_k$ is assumed to be equal to the actual transmitted symbol I_k. For purposes of comparison, the error rate based on exact knowledge of the equalizer coefficients (LMS sense) is also illustrated. Both the performance based on theoretical tap-weights and square root Kalman filters are depicted. In Fig. 4, the performance of the DFE with the square root and revised square root Kalman algorithms exhibit 1.25 and 1 dB degradation respectively related to the ideal performance at $P_e = 10^{-3}$. The performance of the DFE with the revised square root algorithm is slightly better than that with the square root algorithm. This result is due to the improvement of stability of the revised square root Kalman algorithm. In Fig. 5, the revised square root Kalman algorithm has been modified to incorporate the periodic resetting of the U-D factors according to (7.1) and (7.2). The resetting of matrices may slightly degrade the performance in floating-point arithmetic (as indicated in Fig. 5) but will considerably improve the performance in fixed-point arithmetic. The resetting scheme is used to prevent the propagation of roundoff errors associated with the adaptation history.

APPENDIX

The U-D factors of (5.18)

$$\tilde{P} = D'(k-1) - \alpha^{-1}V(k-1)V^{*\prime}(k-1) \tag{A.1}$$

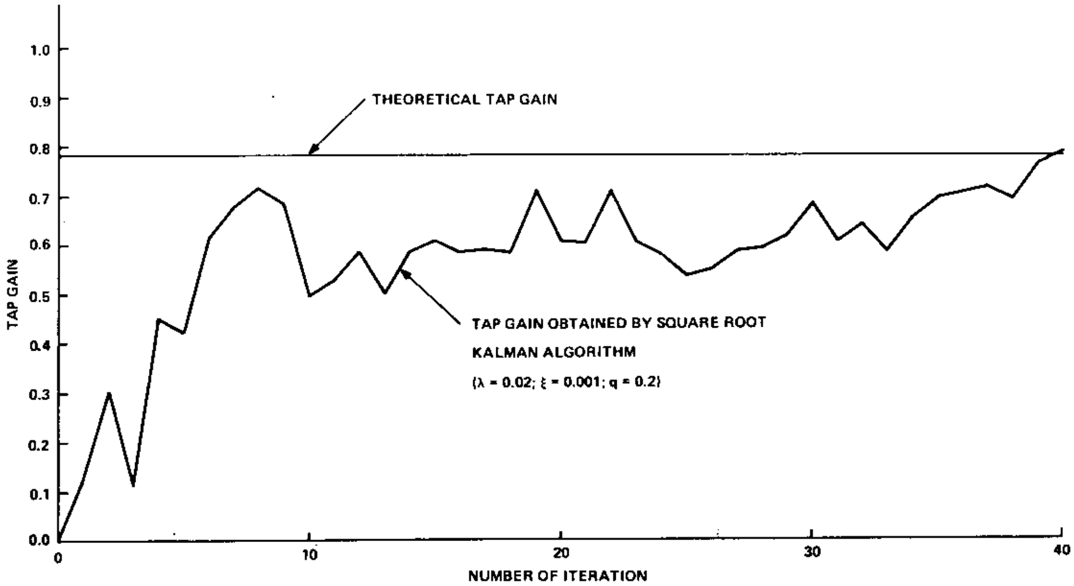

Fig. 1. Properties of convergence of the equalizer tap weight (real part of the main tap shown; DFE; 8ϕ; two-path fading channel; 1 hz fade rate; $N = 5$; SNR = 20 dB).

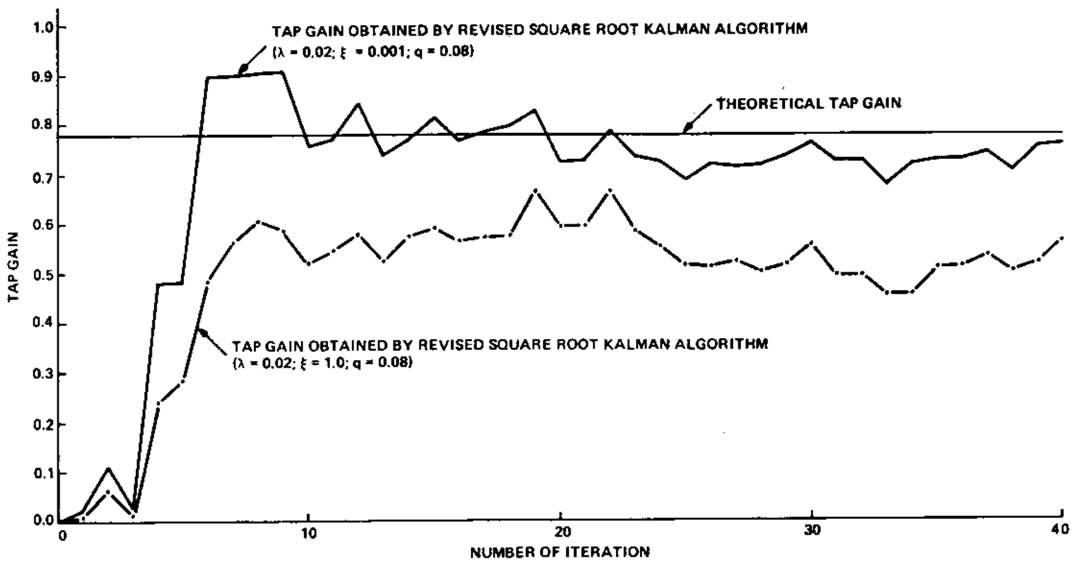

Fig. 2. Properties of convergence of the equalizer tap weight (real part of the main tap shown; DFE; 8ϕ; two-path fading channel; 1 Hz fade rate; $N = 5$; SNR = 20 dB).

may be found by forming the quadratic form $X'\bar{P}X^*$,

$$\begin{aligned} X'\bar{P}X^* &= X'\bar{U}^*\bar{D}\bar{U}'X^* \\ &= \bar{F}^{*\prime}\bar{D}\bar{F} \\ &= \sum_{j=1}^{N} \bar{d}_j |\bar{f}_j|^2, \end{aligned} \tag{A.2}$$

where

$$\bar{F} = \bar{U}'X^*, \tag{A.3}$$

$$\bar{f}_1 = x_1^*, \tag{A.4}$$

$$\bar{f}_j = \sum_{i=1}^{j-1} \bar{\mu}_{ij} x_i^* + x_j^*, \qquad j = 2, 3, \cdots, N. \tag{A.5}$$

Let us suppose $w_N = -\alpha^{-1}$, then (A.2) becomes

$$\begin{aligned} X'\bar{P}X^* &= X'D'X^* + w_N X'VV^{*\prime}X^* \\ &= \sum_{j=1}^{N} d_j' |x_j|^2 + w_N \left| \sum_{j=1}^{N} v_j x_j \right|^2 \\ &= \sum_{j=1}^{N-1} d_j' |x_j|^2 + \left(d_N' + w_N |v_N|^2\right) |x_N|^2 \\ &\quad + w_N \left| \sum_{j=1}^{N-1} v_j x_j \right|^2 + w_N v_N x_N \left(\sum_{j=1}^{N-1} v_j^* x_j^* \right) \\ &\quad + w_N v_N^* x_N^* \left(\sum_{j=1}^{N-1} v_j x_j \right). \end{aligned} \tag{A.6}$$

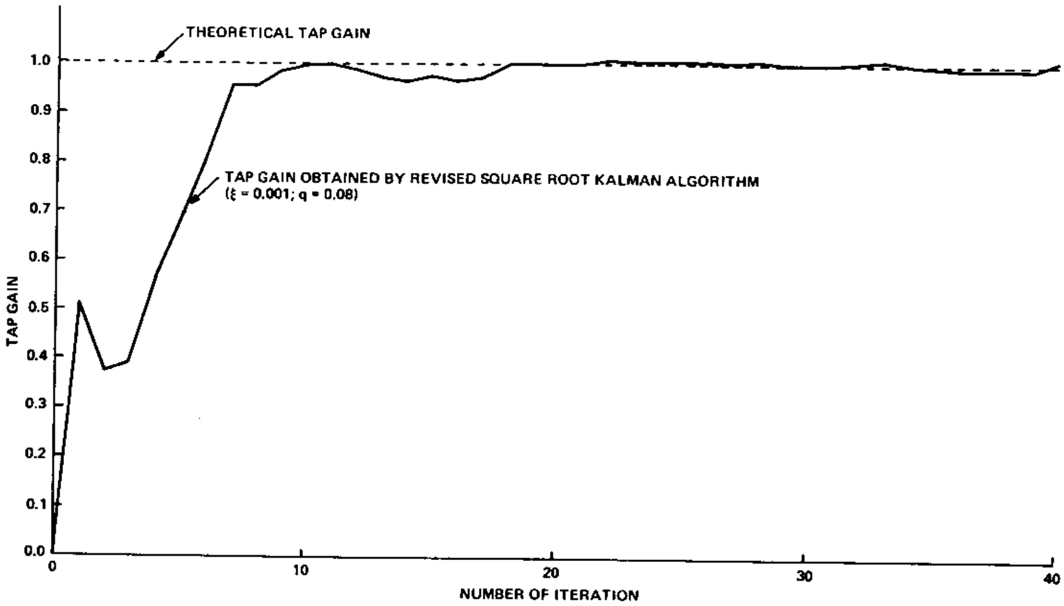

Fig. 3. Properties of Convergence of the Equalizer tap weight (main tap shown; DFE; 8ϕ; two-path fixed channel; $N = 5$; SNR = 30 dB).

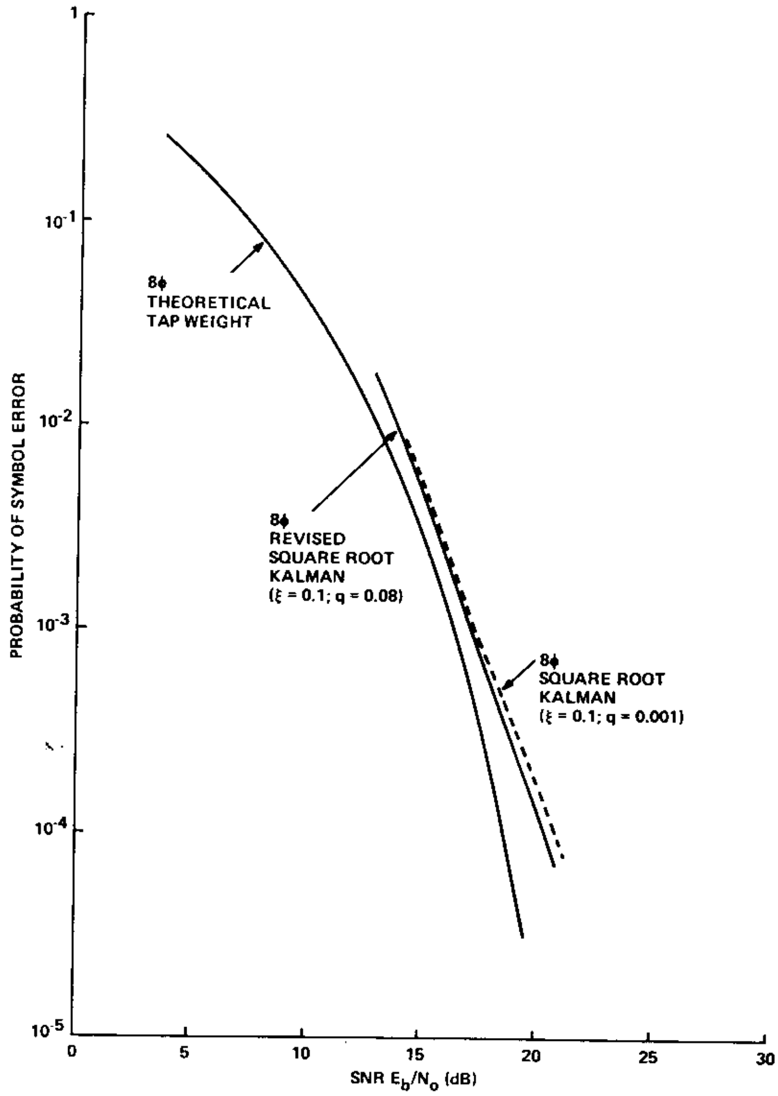

Fig. 4. Error rate performance (DFE; two-path fading channel; 1 Hz fade rate; $\lambda = 0.02$).

provided that

$$\bar{d}_j = d_j' + w_j |v_j|^2, \quad j = 1, 2, \cdots, N, \tag{A.12}$$

$$w_{j-1} = \frac{w_j d_j'}{\bar{d}_j}, \quad j = 1, 2, \cdots, N, \tag{A.13}$$

and

$$\bar{\mu}_{ij} = \frac{w_j v_j}{\bar{d}_j} v_i^*, \quad i = 1, 2, \cdots, j-1; \; j = 2, 3, \cdots, N. \tag{A.14}$$

Multiplying (A.12) by $1/d_j' w_j$, one obtains

$$\frac{1}{w_{j-1}} = \frac{1}{w_j} + \frac{1}{d_j'} |v_j|^2, \quad j = 1, 2, \cdots, N. \tag{A.15}$$

Substituting (A.15) by the following equation

$$d_j' = (1+q) d_j = h_q d_j, \quad j = 1, 2, \cdots, N, \tag{A.16}$$

we have

$$\begin{aligned} \frac{1}{w_{j-1}} &= \frac{1}{w_j} + \frac{d_j}{h_q} |f_j|^2 \\ &= \frac{1}{w_N} + \frac{1}{h_q} \sum_{i=j}^{N} d_i |f_i|^2 \\ &= -\frac{q\alpha + \alpha_{j-1}}{h_q}, \quad j = 1, 2, \cdots, N, \end{aligned} \tag{A.17}$$

where

$$\alpha_j = \sum_{i=1}^{j} d_i |f_i|^2 + \xi, \quad j = 1, 2, \cdots, N, \tag{A.18}$$

$$\alpha = -w_N^{-1} = \alpha_N = \sum_{i=1}^{N} d_i |f_i|^2 + \xi, \tag{A.19}$$

$$\alpha_0 = \xi. \tag{A.20}$$

Substituting (A.16) and (A.17) into (A.13) and (A.14), we obtain

$$\bar{d}_j = d_j (k-1) h_q \frac{\alpha_{j-1} + h_t}{\alpha_j + h_t}, \quad j = 1, 2, \cdots, N, \tag{A.21}$$

$$\bar{\mu}_{ij} = -\frac{h_q v_i^* v_j}{(\alpha_j + h_t) \bar{d}_j}, \quad i = 1, 2, \cdots, j-1; \; j = 2, 3, \cdots, N, \tag{A.22}$$

where

$$h_t = \alpha q. \tag{A.23}$$

Acknowledgment

The author expresses his appreciation to P. Anderson for help and reviews of this work.

References

[1] H. de Pedro, F. M. Hsu, A. A. Giordano, and J. G. Proakis, "Signal design for high speed serial transmission on fading dispersive channels," in *Nat. Telecommun. Conf.*, Birmingham, AL, Dec. 1978.

[2] F. M. Hsu, A. A. Giordano, H. E. de Pedro, and J. G. Proakis, "High speed modem techniques for fading dispersive channels," in *NUSC Workshop Comm. Fading Dispersive Medium*, New London, CT, June 5–6, 1979.

[3] G. J. Bierman, *Factorization Methods for Discrete Sequential Estimation*. New York: Academic, 1977.

[4] D. Falconer and L. Ljung, "Application of fast Kalman estimation to adaptive equalization," *IEEE Trans. Comm.*, Oct. 1978.

[5] T. L. Lim and M. S. Mueller, "Rapid equalizer start-up using least-squares algorithms," in *Int. Conf. Comm.*, 1980.

[6] N. A. Carlson, "Fast triangular formulation of the square root filter," *AIAA J.*, vol. 11, no. 9, Sept., 1973.

[7] N. A. Carlson and A. F. Culmone, "Efficient algorithms for on-board array processing," in *Int. Conf. Commun.*, Boston, MA, June 10–14, 1979.

[8] A. Andrews, "A square root formulation of the Kalman covariance equations," *AIAA J.*, vol. 6, no. 6, June 1968.

[9] J. F. Bellantoni and K. W. Dodge, "A square root formulation of the Kalman-Schmidt filter," *AIAA J.*, vol. 5, no. 7, July 1967.

[10] R. H. Battin, *Astronautical Guidance*. New York: McGraw-Hill, 1964, pp. 338–339.

[11] S. F. Schmidt, "Computational techniques in Kalman filtering," in *Theory and Applications of Kalman Filtering*, C. T. Leondes, Ed., Advisory Group for Aerospace Res. Develop., AGARDOgraph 139, 1970.

[12] G. T. Bierman, "Measurement updating using the *U-D* factorization," in *Proc. IEEE Conf. Decision Contr.*, Houston, TX, 1975.

[13] W. M. Gentleman, "Least squares computations by Givens transformations without square roots," *J. Inst. Math. Appl.* vol. 12, 1973.

[14] D. Godard, "Channel equalization using a Kalman filter for fast data transmission," *IBM J. Res. Develop.*, May 1974.

[15] F. H. Hsu, A. A. Giordano, H. de Pedro and J. G. Proakis, "Adaptive equalization techniques for high-speed transmission on fading dispersive HF channels," in *NTC* '80, Houston, TX, 1980.

[16] E. H. Satorius and J. D. Pack, "Application of least square lattice algorithms to adaptive equalization," *IEEE Commun.*, vol. COM-29, Feb., 1981.

[17] J. G. Proakis, "Advances in equalization for intersymbol interference," in *Advances in Communication Systems, Theory and Application.*, New York: Academic, 1975.

[18] M. J. Shensa, "A least-squares decision feedback equalizer," in *Int. Conf. Commun.*, 1980.

[19] V. Allen and P. Anderson, "Wideband digital modem for HF radio," in *26th Annual Southeastern ISA Conf.*, Orlando, FL, 1980.

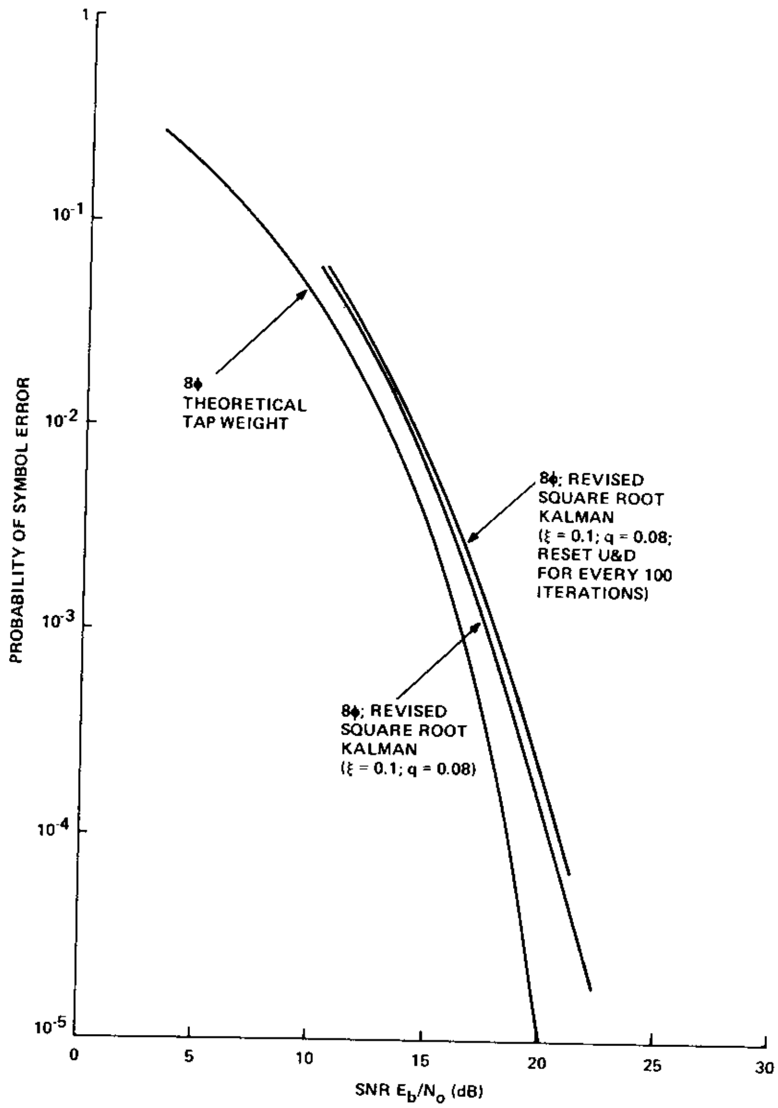

Fig. 5. Error rate performance (DFE; two-path fading channel; 1 Hz fade rate; $\lambda = 0.02$).

Denote the quantities $\bar{d}_N$, W_{N-1}, and $\bar{\mu}_{jN}$ by

$$\bar{d}_N = d'_N + w_N |v_N|^2, \tag{A.7}$$

$$w_{N-1} = \frac{w_N d'_N}{\bar{d}_N}, \tag{A.8}$$

$$\bar{\mu}_{jN} = \frac{w_N v_N}{\bar{d}_N} v_j^*, \qquad j = 1, 2, \cdots, N-1, \tag{A.9}$$

Equation (A.6) may be written as

$$\begin{aligned} X'\bar{P}X^* &= \sum_{j=1}^{N-1} d'_j |x_j|^2 + \bar{d}_N |x_N|^2 + w_N \frac{d'_N + w_N |v_N|^2}{\bar{d}_N} \Big| \sum_{j=1}^{N-1} v_j x_j \Big|^2 + w_N v_N x_N \left(\sum_{j=1}^{N-1} v_j^* x_j^* \right) + w_N v_N^* x_N^* \left(\sum_{j=1}^{N-1} v_j x_j \right) \\ &= \sum_{j=1}^{N-1} d'_j |x_j|^2 + w_{N-1} \Big| \sum_{j=1}^{N-1} v_j x_j \Big|^2 + \left| \sum_{j=1}^{N-1} \bar{\mu}_{jN} x_j^* + x_N^* \right|^2 \bar{d}_N \\ &= \sum_{j=1}^{N-1} d'_j |x_j|^2 + w_{N-1} \left| \sum_{j=1}^{N-1} v_j x_j \right|^2 + \bar{d}_N |\bar{f}_N|^2. \end{aligned} \tag{A.10}$$

Similarly, the inductive reduction follows recursively, and after $N-1$ steps the quadratic form has been reduced to

$$X'\bar{P}X^* = \sum_{j=1}^{N} \bar{d}_j |\bar{f}_j|^2 \tag{A.11}$$