AT&T Technical Journal
Vol. 64, No. 10, December 1985
Printed in U.S.A.

Cross-Polarization Cancellation and Equalization in Digital Transmission Over Dually Polarized Multipath Fading Channels

By M. KAVEHRAD and J. SALZ*

(Manuscript received July 11, 1984)

A theory for data-aided equalization and cancellation in digital data transmission over dually polarized fading radio channels is presented. The present theory generalizes and extends previous work by admitting decision feedback structures with finite-tap transversal filter implementations. Subject to the assumption that some past and/or future data symbols are correctly detected, formulas and algorithms for evaluating the least mean-square error for different structures are presented. In a sequence of curves we evaluate and compare the performance of various structures for a particular propagation model and several fading events. We find improvement in performance for decision feedback over linear equalization. More importantly, we discovered that in this application, as in the single-channel transmission case, decision feedback/canceler structures are much less sensitive to timing phase than linear equalizers.

I. INTRODUCTION

One of the purposes of this article is to call attention to mounting research results pointing the way toward effective methods for combating the deleterious effects of various impairments arising in digital data transmission over dually polarized fading radio channels.

Transmission of M-state Quadrature Amplitude-Modulated (QAM) signals via orthogonally polarized carriers is an effective method for

* Authors are employees of AT&T Bell Laboratories.

reusing existing bandwidth with obvious economic advantages. The main obstacle in the way of realizing these advantages is the unavoidable presence of Cross-Polarization Interference (CPI) between the dually polarized signals that arise due to multipath fading, antenna misalignments, and imperfect waveguide feeds. The chief purpose of our current work is to obtain a fundamental understanding and a solution to this problem.

There is a well established theory of linear and decision feedback equalization/cancellation to mitigate the effects of intersymbol interference (ISI) and noise in the transmission of a single digital signal.[1] However, consideration of data-aided CPI cancellation in addition to ISI equalization in the presence of noise has not been treated before. The work of Amitay and Salz[2] establishes a theoretical base for optimal linear compensation of CPI and ISI in the presence of noise; however, their work is limited strictly to linear techniques and considers only ideal infinite-tap transversal structures.

In this article we generalize previous treatments of this subject in two major respects. Our first contribution is to cast the problem of CPI cancellation and ISI equalization in a general theoretical framework that admits data-aided decision feedback techniques. Secondly, and most importantly, we admit finite-tap transversal structures that in practice can be implemented adaptively.

The receiver configuration is based on a matrix structure suggested by the theory of optimal detection and is shown in Figs. 1 and 2. The optimal structure is comprised of a linear matrix equalizer/canceler and an ISI and CPI estimator, which is used to subtract some of the interference from the received signals. An architecture, previously proposed by Kavehrad,[3] is a special case of this generalized structure.

The dually polarized channel is modeled by a particular 4×4 real matrix impulse response or its Fourier transform followed by additive noise. The 2×2 block-diagonal elements of this matrix represent the copolarized (in line) responses, while the off-diagonal 2×2 block entries represent cross-coupled and cross-polarized interfering responses. Each matrix channel characterizes a snapshot of a multipath fading event, which in the presence of noise limits the achievable error rate of the receiver for a given data rate. We use a propagation model proposed in Ref. 2.

In comparing the performance of various equalizer/cancelers, the Mean-Square Error (MSE) is used. The justification for using this criterion has been amply discussed in the literature.[1,2] But the chief motivation for its use is due to its mathematical tractability. It turns out that it also leads to an exponentially tight upper bound on error rate. In practice, it lends itself to easy estimation and thereby is used to update transversal filter-tap coefficients recursively.

Section II contains the system model and theoretical developments. Computational algorithms are provided in Section III, and our numerical results and associated discussions are given in Section IV. Finally, a summary is presented in the last section.

II. THE MODEL AND THEORETICAL DEVELOPMENTS

2.1 System model

Consider a dually polarized digital radio communications channel supporting two independent QAM data signals. This type of communication channel with an ideal QAM modulator and demodulator is shown in Fig. 1. The four independent synchronous data signals $S_{lv}(t)$, $S_{lh}(t)$, $l = 1, 2$, with the generic representation

$$S_{lv}(t) = \sum_n a_{lvn} g(t - nT), \quad l = 1, 2$$

$$S_{lh}(t) = \sum_n a_{lhn} g(t - nT), \quad l = 1, 2, \tag{1}$$

amplitude modulate two linearly polarized carrier waves in quadrature. The modulated signal,

$$S_v(t) = S_{1v}(t)\cos \omega_0 t + S_{2v}(t)\sin \omega_0 t, \tag{2a}$$

is transmitted over the vertically polarized channel, while

$$S_h(t) = S_{1h}(t)\cos \omega_0 t + S_{2h}(t)\sin \omega_0 t \tag{2b}$$

is transmitted over the horizontal channel. The carrier frequency is ω_0 and the real data symbols

$$\{a_{lvn},\ l = 1, 2\} \quad \text{and} \quad \{a_{lhn},\ l = 1, 2\}, \quad -\infty < n < \infty$$

are assumed to be independently drawn from a lattice of points with odd integer coordinates. The QAM constellations associated with eq. (2) are, therefore, rectangular. The scalar shaping pulse, $g(t)$, is selected by the designer to satisfy limitations on transmitted power and bandwidth.

The individual transmission channels are characterized by bandpass impulse responses or by their respective Fourier transforms,

$$\begin{bmatrix} h_v(t) \\ h_h(t) \end{bmatrix} = \begin{bmatrix} h_{i11}(t) \\ h_{i22}(t) \end{bmatrix} \cos \omega_0 t + \begin{bmatrix} h_{q11}(t) \\ h_{q22}(t) \end{bmatrix} \sin \omega_0 t. \tag{3}$$

The resolution of $h_v(t)$ and $h_h(t)$ into their respective baseband in-phase and quadrature components turns out to be convenient in our application.

To accommodate coupling between the polarized channels, two pairs of impulse responses, one associated with the cochannel and the other

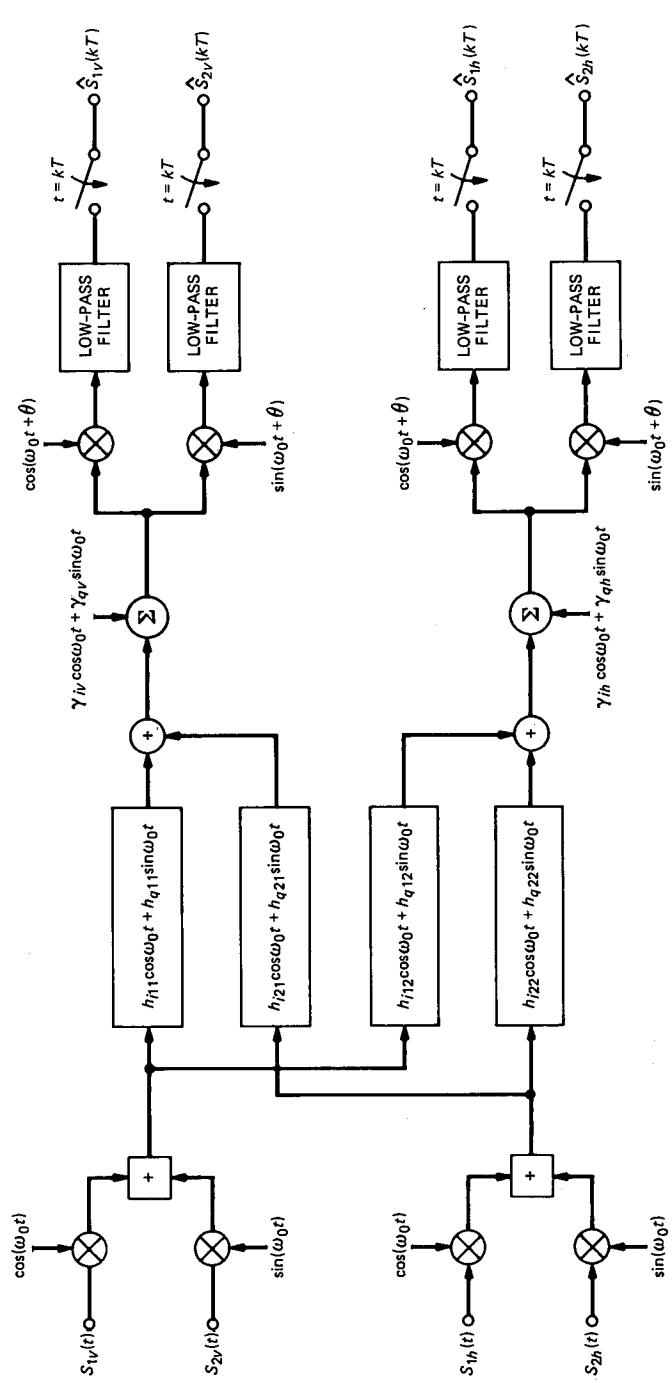

Fig. 1—System block diagram.

associated with the cross-channel, are used to completely characterize the medium.

At the output, two independent noises are added and the signal plus noise is then coherently demodulated. The end-to-end system including the modulators and the demodulators is shown in Fig. 1. It is convenient to view this linear system as a four-input port four-output port network and characterize it by a 4×4 matrix impulse response or its Fourier transform, which is the overall system frequency response.

It is now easy to verify that the I/O relationships can be expressed as follows (see Fig. 1):

$$\begin{aligned}
D_{1v} &= S_{1v}*h_{i11} + S_{2v}*h_{q11} + S_{1h}*h_{i21} + S_{2h}*h_{q21} + \nu_{iv} \\
D_{2v} &= -S_{1v}*h_{q11} + S_{2v}*h_{i11} - S_{1h}*h_{q21} + S_{2h}*h_{i21} + \nu_{qv} \\
D_{1h} &= S_{1v}*h_{i12} + S_{2v}*h_{q12} + S_{1h}*h_{i22} + S_{2h}*h_{q22} + \nu_{ih} \\
D_{2h} &= -S_{1v}*h_{q12} + S_{2v}*h_{i12} - S_{1h}*h_{q22} + S_{2h}*h_{i22} + \nu_{qh}, \qquad (4)
\end{aligned}$$

where $*$ denotes convolution,

$$h*S = \int_{-\infty}^{\infty} h(t-\tau)S(\tau)d\tau.$$

The representation in eq. (4) can be put into a convenient matrix form,

$$D(t) = \int_{-\infty}^{\infty} H(t-\tau)S(\tau)d\tau + \nu(t), \qquad (5)$$

where $H(t)$ is the 4×4 matrix channel impulse response

$$H(t) = \begin{bmatrix} h_{i11}(t) & h_{q11}(t) & h_{i21}(t) & h_{q21}(t) \\ -h_{q11}(t) & h_{i11}(t) & -h_{q21}(t) & h_{i21}(t) \\ h_{i12}(t) & h_{q12}(t) & h_{i22}(t) & h_{q22}(t) \\ -h_{q12}(t) & h_{i12}(t) & -h_{q22}(t) & h_{i22}(t) \end{bmatrix}, \qquad (6)$$

$$S(t) = \begin{bmatrix} S_{1v}(t) \\ S_{2v}(t) \\ S_{1h}(t) \\ S_{2h}(t) \end{bmatrix} \qquad (7)$$

is the input signal vector, and

$$\nu(t) = \begin{bmatrix} \nu_{iv}(t) \\ \nu_{qv}(t) \\ \nu_{ih}(t) \\ \nu_{qh}(t) \end{bmatrix} \qquad (8)$$

is the added noise vector.

Since complex numbers $x + jy$ are isomorphic to matrices of the form

$$x + jy \sim \begin{bmatrix} x & y \\ -y & x \end{bmatrix},$$

the channel model as described by the 4×4 real matrix, eq. (6), can also be represented by a 2×2 complex matrix of the form[2]

$$\begin{bmatrix} h_{i11} + jh_{q11} & h_{i21} + jh_{q21} \\ h_{i12} + jh_{q12} & h_{i22} + jh_{q22} \end{bmatrix}.$$

In our application, however, it turns out to be more convenient to work with the real matrix in eq. (6).

We now return to the I/O relationship in eq. (5) and substitute eq. (1) to obtain in more detail

$$D(t) = \sum_n \int g(\tau - nT)H(t - \tau)d\tau \cdot A_n + \nu(t), \tag{9}$$

where the real data symbol vector A_n is given by

$$A_n = \begin{bmatrix} a_{1vn} \\ a_{2vn} \\ a_{1hn} \\ a_{2hn} \end{bmatrix}. \tag{10}$$

A representative sample of $D(t)$ taken at $t = 0$, without loss of generality, yields

$$D(0) = H_0A_0 + \sum_{\substack{n \\ n \neq 0}} H_nA_n + \nu(0), \tag{11}$$

where

$$H_n = \int_{-\infty}^{\infty} g(\tau - nT) \times H(-\tau)d\tau. \tag{12}$$

In an ideal system, eq. (11) would yield $D(0) = A_0 \times$ constant. This result is obtained when

1. $H_0 =$ constant $\times I$ (I is the identity matrix), which implies that the flat, or nondispersive, CPI vanishes;
2. $H_n = [0]$, ([0] is the zero matrix), implying that CPI, as well as ISI, vanishes; and
3. $\nu(0) = 0$.

Clearly, these requirements cannot be achieved in practice, and the designer of data communications systems must deal with these impairments and find methods that minimize their effects on system performance.

A well-known approach[2] is the use of linear equalization. Our objective here is to investigate a general cancellation technique in conjunction with linear equalization, which could potentially yield better performance than with just the linear equalizer alone. To this end, we begin our analysis by placing a linear matrix filter in cascade with the channel prior to sampling, and we choose its characteristics so as to minimize the total MSE between the actual output sample and the desired output after canceling some CPI and ISI.

Denote the matrix filter impulse response by $W(t)$ and evaluate its output at $t = 0$. This yields the column vector for the overall system response

$$D_0(0) = U_0A_0 + \sum_{\substack{n \\ n\neq 0}} U_nA_n + \nu_0, \tag{13}$$

where

$$U_n = \int_{-\infty}^{\infty} W(-\tau)H_0(\tau - nT)d\tau,$$

$$H_0(t) = \int_{-\infty}^{\infty} g(\tau)H(t - \tau)d\tau, \tag{14}$$

and

$$\nu_0 = \int_{-\infty}^{\infty} W(-\tau)\nu(\tau)d\tau. \tag{15}$$

2.2 The optimization problem

To describe our approach, we first discuss the following statistical problem. Suppose that one observes the vector $D_0(0)$, eq. (13), and wishes to design the best processing strategy that estimates A_0 in a sense of minimizing the probability of error. The precise solution to this problem remains intractable because of the non-Gaussian nature of ISI and CPI. While the precise mathematical solution is unknown, some qualitative aspects of the solution have been discussed.[4,5] It is easy to argue that the optimal detector structure consists of a matched filter followed by a least-mean-square estimator of the interference, which is then subtracted from the matched filter output. After subtracting the estimate of the interference, the problem reduces to detecting a known signal in additive Gaussian noise, which has a well-known solution. The difficulty with this formulation, while physically appealing, is that the least-mean-square estimator of interference is just as difficult and intractable to evaluate as the detection problem originally posed. One redeeming feature of this approach, however, is that if one does not insist on least-mean-square estimation of inter-

ference, a reasonable detector structure can be determined. We argue that constructing reasonable estimates of CPI and ISI, which are not necessarily optimum, subtracting them from the incoming signal, and then constructing an optimum detector essentially satisfies the spirit of the suggested optimal procedure.

We now formulate our approach more precisely. To start, assume that over a finite set of sampling instants, S, vector data symbols, A_n, $n \in S$, are available at the receiver and before we make a final decision on the current symbol, A_0, a portion of the interference,

$$\sum_{n \in S} U_n A_n,$$

is subtracted from $D_0(0)$. Actually, this is feasible since prior to $n = 0$, symbols have been decoded all along and what is presumed in our proposal is that we use the already-decoded symbols to improve on the current estimate of A_0. Since practical systems are not realizable relative to a large delay, there is a problem in using symbols that have not yet occurred. This can be overcome by introducing a delay, making tentative decisions, and then returning to modify the A_0 decision.

How realistic is this assumption? The answer depends on the system error rate prior to cancellation. For example, when the error rate is 10^{-4} and the cancellation window size is small relative to 10^4, the probability that almost all of the symbols in this window have been correctly detected is fairly large. Thus, after cancellation, the error rate may be much improved. On the other hand, if the error rate prior to cancellation is high, no improvement after cancellation can be expected since the estimation of the interference is not reliable. Evidently, decision-directed cancellation as proposed here is a bootstrapping technique. It is very successful over a certain range of error rates and fails when the error rate is high. Unfortunately, these qualitative statements are extremely difficult to make precise, and it is necessary to rely on simulation results.[6] The assumption that A_n is known in the canceler window will clearly result in optimistic performance predictions, and whether the predicted benefits can be realized must be ascertained experimentally.

We now proceed to include this "genie" in our mathematical analysis. As already stated, the performance criterion we use throughout this work is the least MSE normalized to the transmitted symbols variance, denoted σ_d^2. This is a mathematically tractable criterion to work with, and by minimizing MSE, one also minimizes an exponentially tight upper bound on the error rate. Its use is also practically motivated because it lends itself to easy estimation, and it can be used to update transversal filter-tap coefficients in practical adaptive systems.

Returning to the mathematical problem at hand, we define the error vector ϵ as the difference between $D_0(0)$ minus the canceler output vector, and the desired vector data symbol, A_0,

$$\epsilon = U_0 A_0 + \sum_{\substack{n \\ n \neq 0}} U_n A_n - \sum_{n \in S} C_n A_n + \nu_0 - A_0, \tag{16}$$

where C_n represents canceler-tap values. Total MSE can be expressed as

$$\text{MSE} = \text{tr}[E\{\epsilon\epsilon^\dagger\}], \tag{17}$$

where "tr" stands for trace of a matrix, $E\{\cdot\}$ denotes mathematical expectation with respect to all random variables, and † represents complex conjugate transpose.

The computation of eq. (17) is straightforward and yields

$$\text{MSE} = \sigma_d^2 \text{tr}\left[I - U_0 - U_0^\dagger + \sigma^2 \int_{-\infty}^{\infty} W(t)W^\dagger(t)dt + \sum_{n \in S} (U_n - C_n)(U_n - C_n)^\dagger + \sum_{n \notin S} U_n U_n^\dagger\right], \tag{18}$$

where $I\sigma_d^2 = E\{A_n A_n^\dagger\}$, $\sigma_d^2 = 2(M-1)/3$, and M is the total number of QAM signal states,

$$N_0 I = E\{\nu(t)\nu^\dagger(t)\},$$

and $\sigma^2 = N_0/\sigma_d^2$.

The set of canceler matrices, C_n, $n \in S$, can immediately be determined. If they are not identically set to U_n, they can only increase the value of MSE. Consequently, we set $C_n = U_n$, $n \in S$, and the residual MSE results in a functional of the matrix impulse response, $W(t)$, and the size of the cancellation window.

The minimization of MSE with respect to the matrix $W(t)$ is accomplished by the use of the calculus of variations. After substituting for U_n, defined in eq. (14), we get

$$\frac{\text{MSE}}{\sigma_d^2} = \text{tr}\left[I - 2\int_{-\infty}^{\infty} W(-\tau)H_0(\tau)d\tau + \sigma^2 \int_{-\infty}^{\infty} W(-\tau)W^\dagger(-\tau)d\tau + \sum_{n \notin S} \int W(-\tau)H_0(\tau - nT)d\tau \int H_0^\dagger(\tau - nT)W^\dagger(-\tau)d\tau\right]. \tag{19}$$

To determine the optimum W, we replace the matrix W in eq. (19) by

$$(W_0)_{ij} + (\xi\eta)_{ij}, \quad i, j = 1, \cdots, 4,$$

where η_{ij} is arbitrary, and we set

$$\frac{\partial}{\partial \xi_{ij}} (\text{MSE}) = [0]_{ij} \tag{20}$$

at $\xi_{ij} = 0$, $i, j = 1, 2, 3, 4$. It is easy to verify that

$$\frac{1}{\sigma_d^2} \frac{\partial}{\partial \xi_{ij}} (\text{MSE}) = \text{tr} \Bigg[-2 \int_{-\infty}^{\infty} \eta_{ij}^0(\tau) H_0(\tau) d\tau + 2\sigma^2 \int_{-\infty}^{\infty} W_0(-\tau) \eta_{ji}^0(\tau) d\tau + 2 \sum_{n \notin S} \int_{-\infty}^{\infty} W_0(\tau) H_0(\tau - nT) d\tau \int_{-\infty}^{\infty} H_0^\dagger(\tau - nT) \eta_{ji}^0(\tau) d\tau \Bigg] = 0, \quad (i, j) = 1, \cdots, 4, \tag{21}$$

where the matrices, η_{ij}^0, $i, j = 1, \cdots, 4$ have the entry "l" in the (l_{ij})th position and zero everywhere else. By computing the trace of eq. (21), we obtain

$$- \int_{-\infty}^{\infty} [H_0(\tau)]_{ji} \eta_{ij}^0(\tau) d\tau + \sigma^2 \int_{-\infty}^{\infty} [W_0(-\tau)]_{ij} \eta_{ij}^0(\tau) d\tau + \sum_{n \notin S} \int [H_0(\tau - nT) U_n^\dagger]_{ji} \eta_{ij}^0(\tau) d\tau = 0, \quad i, j = 1, \cdots, 4. \tag{22}$$

Since eq. (22) must hold for all functions of τ and $\eta_{ij}^0(\tau)$, we obtain the matrix integral equation that must be satisfied by the optimum matrix $W_0(\tau)$, namely,

$$\sigma^2 W_0(-\tau) = H^\dagger(\tau) - \sum_{n \notin S} U_n H_0^\dagger(\tau - nT). \tag{23}$$

The structure of $W_0(\tau)$ is practically interesting. It consists of a matched filter followed by a matrix-tapped delay line where the matrix taps are zero for $n \in S$. In other words, the linear transversal filter or equalizer specified in eq. (23) operates over a range of matrix-tap coefficients where the canceler is not operative. This is to avoid interaction between collocated taps and possible unstability problems. The structure is shown schematically in Fig. 2. In practice, this structure can be approximated and implemented by a finite transversal filter whose taps can be adaptively updated.

After post-multiplying eq. (23) by $W^\dagger(-\tau)$, integrating, and then comparing the result with eq. (19), we get an explicit formula for the optimum MSE,

$$\text{MSE}_0 = \sigma_d^2 \text{tr}(I - U_0), \tag{24}$$

where U_0 is obtained by solving a set of infinite linear equations obtained by post-multiplying eq. (23) by $H(\tau - kT)$ and then integrating. Thus,

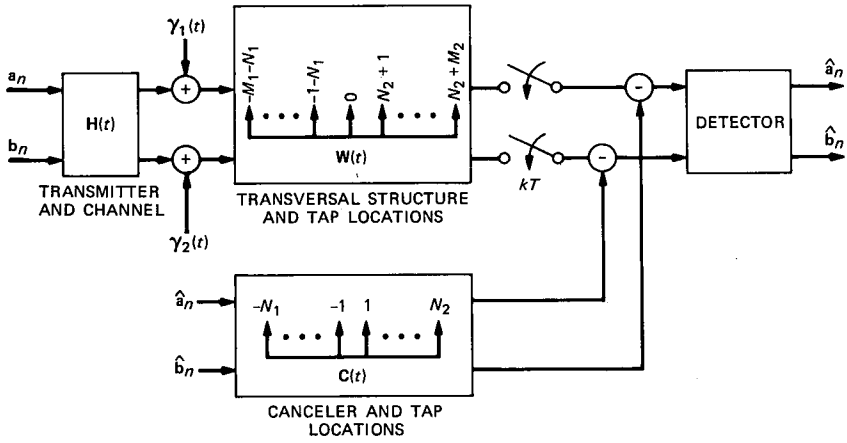

Fig. 2—Cross-polarization interference and intersymbol interference canceler block diagram.

$$\sigma^2 U_k = R_k - \sum_{n \notin S} U_n R_{k-n}, \quad \text{all } k, \tag{25}$$

where

$$R_k = \int_{-\infty}^{\infty} H^\dagger(\tau) H(\tau - kT) d\tau$$
$$= R_{-k}^\dagger. \tag{26}$$

To evaluate the merits of our system, we must have a solution for U_0. The task of solving eq. (25) is rather complicated. It is made difficult by the fact that the matrix equations are not specified over the finite set, S. While the number of unknowns is infinite, the values at the gap window are not specified. A way around this dilemma was found in the scalar case,[4] and with care applied to matrix manipulations, it is possible to adopt the same techniques here.

We proceed by first separating eq. (25) into two equations, one for $k = 0$ and the other for $k \neq 0$. Thus,

$$U_0(I\sigma^2 + R_0) = R_0 - \sum_{n \notin J} U_n R_{-n}, \quad k = 0, \tag{27}$$

and

$$\sum_{n \notin J} U_n M_{k-n} = (I - U_0) R_k, \quad k \notin J, \tag{28}$$

where the set J is defined as

$$\{J: n \in J, \quad n = -N_1, \cdots, 0, \cdots, N_2\} \tag{29}$$

and

$$M_k = R_k + \sigma^2 \delta_{0k} I, \tag{30}$$

where δ_{0k} is the Kronecker delta function. The solution of eq. (28) is facilitated by introducing a set of matrix variables $\{V_n\}_{-\infty}^{\infty}$ and a set of unknown matrices $\{\Lambda_k\}_{-\infty}^{\infty}$. Using these matrices, we write eq. (28) as

$$\sum_{n=-\infty}^{\infty} V_n M_{k-n} = (I - U_0)(R_k - \Lambda_k), \quad \text{all } k. \tag{31}$$

For these doubly infinite sets of matrix equations to identically coincide with eq. (28), the following constraints must hold:

$$\Lambda_n = 0, \quad n \notin J$$

and

$$V_n = 0, \quad n \in J. \tag{32}$$

If these can be satisfied, the solution to eq. (31) will be identical to the solution of eq. (28) with $V_n = U_n$, $n \notin J$, and this is the sole purpose for introducing new variables. Evidently, eq. (31) is easy to solve since it is in a form of a convolutional equation. To this end define the inverse matrix sequences, $\{M_n^{(-1)}\}_{-\infty}^{\infty}$, as

$$\sum_{n=-\infty}^{\infty} M_{k-n} M_n^{(-1)} = I\delta_{k0}, \quad \text{all } k. \tag{33}$$

Now, insert this into eq. (31) to obtain explicitly the desired solution,

$$V_n = (I - U_0) \sum_{k=-\infty}^{\infty} (R_k - \Lambda_k) M_{k-n}^{(-1)}, \quad \text{all } n. \tag{34}$$

From this we can obtain a finite set of equations in the unknown matrices Λ_k, since $V_n = 0$ for $n \in J$,

$$\sum_{k=-\infty}^{\infty} R_k M_{k-n}^{(-1)} = \sum_{k \in J} \Lambda_k M_{k-n}^{(-1)}, \quad n \in J. \tag{35}$$

By substituting the definition of M_k from eq. (30) into the left-hand side of eq. (35) and making use of eq. (33), we obtain the desired equations for the unknown constraint matrices Λ_k, $k \in J$,

$$I\delta_{n0} - \sigma^2 M_{-n}^{(-1)} = \sum_{k \in J} \Lambda_k M_{k-n}^{(-1)}, \quad n \in J. \tag{36}$$

Returning to eq. (31), we get for $k = 0$

$$\begin{aligned} \sum_{n=-\infty}^{\infty} V_n M_{-n} &= (I - U_0)(R_0 - \Lambda_0) \\ &= \sum_{n \notin J} U_n R_{-n}, \end{aligned} \tag{37}$$

where the last equality derives from the fact that $V_n = 0$, $n \in J$; $V_n = U_n$, $n \notin J$; and $R_n = M_n$, $n \in J$. Finally, by substituting eq. (27) into eq. (37), we can write

$$(I - U_0)(R_0 - \Lambda_0) = R_0 - U_0(I\sigma^2 + R_0), \tag{38}$$

and solving for $I - U_0$ yields

$$(I - U_0) = \sigma^2(I\sigma^2 + \Lambda_0)^{-1}. \tag{39}$$

Substituting this into eq. (24) provides an explicit expression for MSE_0 in terms of Λ_0 only,

$$\mathrm{MSE}_0 = \sigma_d^2 \mathrm{tr}\left(I + \frac{\Lambda_0}{\sigma^2}\right)^{-1}. \tag{40}$$

Our effort in the following will be centered on determining Λ_0 as a function of the cancellation window size, or the size of set J.

2.3 The matched filter bound

When the canceler window is doubly infinite in extent, one obtains the very best possible result. In other words, the genie has eliminated all ISI and CPI. In this special case, $N_1 = -\infty$ and $N_2 = \infty$, and eq. (36) is now easy to solve since it reads

$$I\delta_{n0} - \sigma^2 M_{-n}^{(-1)} = \sum_{k=-\infty}^{\infty} \Lambda_k M_{k-n}^{(-1)}. \tag{41}$$

By evaluating the Fourier series of both sides of eq. (41), we obtain

$$I = \sigma^2 M^{(-1)}(\theta) + \Lambda(\theta) M^{(-1)}(\theta), \tag{42}$$

where a generic Fourier series pair representation is

$$X(\theta) = \sum_{l=-\infty}^{\infty} x_l \exp(j\theta l)$$

and

$$x_l = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\theta) \exp(-j\theta l) d\theta.$$

Since $M(\theta) = \sigma^2 I + R(\theta)$ and $M^{(-1)}(\theta)$ is in fact the inverse, $M^{-1}(\theta)$, we determine from eq. (42) that $\Lambda(\theta) = R(\theta)$. Consequently, the zeroth coefficient of $\Lambda(\theta)$ is $R_0 = 1/(2\pi) \int_{-\pi}^{\pi} R(\theta) d\theta$, and when this is substituted into eq. (40) we get the desired matched filter bound,

$$\mathrm{MSE}_0 = \sigma_d^2 \mathrm{tr}\left(I + \frac{R_0}{\sigma^2}\right)^{-1}. \tag{43}$$

This will serve as a lower bound to attainable performance to which we will compare all other results.

2.4 Linear equalization

In this case, the canceler is absent and so $N_1 = N_2 = 0$. Here, eq. (36) reduces to

$$I - \sigma^2 M_0^{(-1)} = \Lambda_0 M_0^{(-1)}, \tag{44}$$

and solving for $M_0^{(-1)}$, we get

$$\begin{aligned} M_0^{(-1)} &= \frac{1}{\sigma^2}\left(I + \frac{\Lambda_0}{\sigma^2}\right)^{-1} \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} M^{-1}(\theta)d\theta = \frac{1}{2\pi\sigma^2}\int_{-\pi}^{\pi}\left[I + \frac{R(\theta)}{\sigma^2}\right]^{-1} d\theta. \end{aligned} \tag{45}$$

It is now immediate that

$$\mathrm{MSE}_0 = \frac{\sigma_a^2}{2\pi}\int_{-\pi}^{\pi} \mathrm{tr}\left(I + \frac{R(\theta)}{\sigma^2}\right)^{-1} d\theta, \tag{46}$$

the well-known formula for linear equalization.[2]

2.5 Decision feedback and finite causal canceler

In this application it is assumed that all the causal terms, which depend only on past decisions, are canceled in addition to a finite number of noncausal terms. This implies that $N_2 = \infty$ and N_1 is finite. When $N_1 = 0$, the canceler becomes a decision feedback equalizer[7] since causal interference can be canceled by a feedback circuit. Here, we will determine MSE for the more general case when N_1 is not necessarily zero.

To treat this case it is more convenient to solve for U_0 directly from eq. (28) rather than through eq. (36). Thus, we rewrite eq. (28) as

$$\sum_{k=-\infty}^{-N_1} U_k M_{m-k} = (I - U_0)R_m, \quad m \le -N_1, \tag{47}$$

which is recognized to be a matrix Wiener-Hopf equation, and its solution depends on being able to factor positive definite Hermitian matrices.[8]

To proceed with the solution of eq. (47), we introduce the following sequence of matrices

$$M_n^+ = [0], \quad n < 0$$

$$M_n^- = [0], \quad n \ge 0,$$

such that

$$M_m = \sum_{n=0}^{\infty} M_{m-n}^- M_n^+, \quad \text{all } m. \tag{48}$$

The validity of this expression and the existence of M_n^+ and M_n^- were first proved by Wiener and Akutowicz.[9]

Substituting eq. (48) into eq. (47) and rearranging yields two sets of equations

$$\sum_{n=0}^{\infty} y_{m-n} M_n^+ = (I - U_0) R_m, \quad \text{all } m \tag{49}$$

and

$$\sum_{k=-\infty}^{-N_1} U_k M_{m-k}^- = y_m, \quad m \leq -N_1. \tag{50}$$

The procedure for solving these is to first solve for $Y(\theta)$ from eq. (49) in terms of $M^-(\theta)$, an easy task in terms of the Fourier transforms of $\{M_n^-\}$ and $\{y_n\}$. Having obtained $Y(\theta)$, one proceeds to solve eq. (50) for $U(\theta)$ in terms of $M^\dagger(\theta)$. Note that eq. (48) implies

$$M(\theta) = M^-(\theta) M^+(\theta),$$

and since $M(\theta)$ is Hermitian, $M(\theta) = M^\dagger(\theta)$, implying $[M^+(\theta)]^\dagger = M^-(\theta)$, $[M^-(\theta)]^\dagger = M^+(\theta)$, and the factorization problem is reduced to finding a matrix $M^+(\theta)$ such that

$$M(\theta) = [M^+(\theta)]^\dagger M^+(\theta),$$

where the entries in $M^+(\theta)$, $[M^+(\theta)]_{ij}$ are such that $[M^+(\theta)]_{ij}$ has a Fourier series with only positive frequency coefficients. We shall later discuss algorithms for determining $M^+(\theta)$ from $M(\theta)$—a rather complicated task.[10]

We now proceed to determine the sequence y_m. Multiply both sides of eq. (50) by $M_{-m}^\pm$ and sum m from $-\infty$ to $-N_1$. This gives the formula

$$\sum_{m=-\infty}^{-N_1} y_m M_{-m}^+ = \sum_{k=-\infty}^{-N_1} U_k M_{-k}. \tag{51}$$

Now, recall that $M_k = R_k + \sigma^2 \delta_{k0} I$, and, therefore, eq. (49) can be put into the form

$$\sum_{n=0}^{\infty} y_{m-n} M_n^+ = (I - U_0) R_m, \quad \text{all } m, \quad m \neq 0. \tag{52}$$

When this is compared with eq. (48), we obtain

$$y_m = (I - U_0) M_m^-, \quad m \neq 0, \tag{53}$$

and when substituted into eq. (51), we get

$$(I - U_0) \sum_{m=-\infty}^{-N_1-1} M_m^- M_{-m}^+ = \sum_{m=-\infty}^{-N_1-1} U_m M_{-m} = (I - U_0) \sum_{m=-\infty}^{-N_1-1} M_m^- (M_m^-)^\dagger. \tag{54}$$

From eq. (37) we have that

$$\sum_{n=-\infty}^{-N_1-1} U_n M_{-n} = (I - U_0)(R_0 - \Lambda_0), \tag{55}$$

and so we conclude that

$$R_0 - \Lambda_0 = \sum_{n=-\infty}^{-N_1-1} M_n^- (M_n^-)^\dagger. \tag{56}$$

Substituting again for $R_0 = M_0 - I\sigma^2$ in eq. (56) and rearranging, we finally obtain

$$I\sigma^2 + \Lambda_0 = \sum_{n=-N_1}^{0} M_n^- (M_n^-)^\dagger, \tag{57}$$

since

$$\begin{aligned} M_0 &= \sum_{n=0}^{\infty} M_{-n}^- M_n^+ \\ &= \sum_{n=0}^{\infty} M_{-n}^- (M_{-n}^-)^\dagger \\ &= \sum_{n=-N_1}^{0} M_n^- (M_n^-)^\dagger; \end{aligned}$$

hence,

$$\begin{aligned} I\sigma^2 + \Lambda_0 &= -\sum_{n=-\infty}^{-N_1-1} M_n^- (M_n^-)^\dagger + \sum_{n=-\infty}^{0} (M_n^-)(M_n^-)^\dagger \\ &= \sum_{n=-N_1}^{0} M_n^- (M_n^-)^\dagger. \end{aligned}$$

Upon substituting eq. (57) into eq. (40), we obtain the desired result,

$$\begin{aligned} \mathrm{MSE}_0 &= \sigma_d^2 \mathrm{tr} \left(I + \frac{\Lambda_0}{\sigma^2} \right)^{-1} \\ &= \sigma_d^2 \mathrm{tr} \left(\frac{1}{\sigma^2} \sum_{n=-N_1}^{0} M_n^- (M_n^-)^\dagger \right)^{-1}. \end{aligned} \tag{58a}$$

Notice that when $N_1 = 0$, that is, no anticausal cancellation,

$$\mathrm{MSE}_0 = \sigma_d^2 \mathrm{tr} \left[\frac{M_o^2}{\sigma^2} \right]^{-1}, \tag{58b}$$

where $M_0 = M_0^- = (M_0^+)^\dagger$. This is the formula for decision feedback equalization derived by Falconer and Foschini for QAM transmission over a single channel, which they cast in a matrix formulation.[11] Evidently, the form of the answer generalizes to arbitrary dimensions.

2.6 *Finite linear equalizer*

The theoretical results we have derived so far apply to an ideal canceler of any window size and an infinite-tap linear equalizer whose matrix taps vanish inside the cancellation window. To assess the penalties incurred by a finite-tap linear equalizer outside the cancellation window, we derive least MSE formulas applicable to this case. With these formulas we will be in a position to evaluate the merits of equalization/cancellation using only a finite number of matrix taps and to gain insight as to how best to deploy the total number of available taps. Also inherent in the theory derived so far is the independence of MSE on sampling phase. This is so since the transversal equalizer/canceler is preceded by a matched filter whose structure presumes knowledge of sampling phase. Here, we shall relax this condition and derive the MSE for a front-end filter matched to the transmitter filter only rather than to the overall channel response and, thereby, bring out the dependence of MSE on timing phase.

We, thus, represent the finite-tap delay line matrix filter by

$$W_{f0}(-\tau) = \sum_{n\in F} g(t - nT)Q_n, \tag{59}$$

where the two sets F and S are disjoint and F now is a finite set,

$$\{F\colon n \in F,\ n = -N_1 - M_1, \cdots, -N_1 - 1, 0, N_2 + 1, \cdots, N_2 + M_2\}.$$

In eq. (59), $g(t)$, as before, is a scaler pulse shape, while $\{Q_n\}_{n\in F}$ is a 4×4 matrix sequence. The objective now is to select the Q_n's that minimize the total MSE, eq. (19),

$$\frac{\text{MSE}}{\sigma_d^2} = \text{tr}\left[I - 2U_0 + \sigma^2 \int_{-\infty}^{\infty} W_f(\tau) W_f^\dagger(\tau) d\tau + \sum_{n\notin S} U_n U_n^\dagger\right]. \tag{60}$$

Substituting eq. (59) into eq. (60) yields

$$\frac{\text{MSE}}{\sigma_d^2} = 2 - 2 \sum_{n\in F} \text{tr}(Q_n H_{-n}) + \sum_{n,m\in F} \text{tr}(Q_n G_{nm} Q_m^\dagger) + \sigma^2 \sum_{n,m\in F} \text{tr}(Q_n \rho_{n-m} Q_m^\dagger), \tag{61}$$

where the H_n's are defined in eq. (12) and

$$\begin{cases} G_{nm} = \sum_{l\notin S} H_{l-n} H_{l-m}^\dagger \\ \rho_n = \int_{-\infty}^{\infty} g(t) g(t - nT) dt. \end{cases} \tag{62}$$

Setting the derivatives of eq. (61) with respect to the elements of the matrices $\{Q_n\}_{n\in F}$ to zero, we get a set of linear matrix equations for the unknowns, $\{Q_n\}_{n\in F}$, namely,

$$H_{-n}^{\dagger} = \sum_{l \in F} Q_l R_{ln}, \quad n \in F, \tag{63}$$

where

$$R_{ln} = G_{ln} + \sigma^2 \rho_{n-l}, \quad n, l \in F. \tag{64}$$

The solution of eq. (63) is straightforward and is discussed in a later section.

For now, label the solution of eq. (63) by Q_l^0—the optimal Q_l's. Premultiply by Q_n^0, sum over $n \in F$, and substitute the result into eq. (63). This yields the desired formula for the least-mean-square error,

$$\mathrm{MSE}_0 = \sigma_d^2 \mathrm{tr}\left(I - \sum_{l \in F} Q_l^0 H_{-l}\right). \tag{65}$$

The next section will present computation algorithms for numerically evaluating the formulas developed here.

III. COMPUTATIONAL ALGORITHMS

An examination of Section III demonstrates that the theoretical analysis of M-QAM signal transmission over dually polarized channels in the presence of multipath fading is a numerically intensive activity. In this section we provide an overview of the major computational issues related to our investigation.

3.1 Infinite linear equalizer/finite canceler

When the linear equalizer in Fig. 2 has a finite-tap window size, the optimum receiver structure comprises a matched filter followed by a matrix transversal filter and a matrix canceler. The most general case under this assumption is when the matrix canceler has a finite number of causal and anticausal taps and the solution of eq. (36) for Λ_0 provides a means of calculating minimum mean-square error by use of eq. (40). To solve for Λ_0, block matrices $M_k^{(-1)}$'s defined in eq. (33) have to be determined first. One way to determine the $M_k^{(-1)}$'s is to solve eq. (33) by a Levinson-type algorithm[12] where the entries are block matrices. Thus, matrix convolution eq. (33) is then represented as

$$\begin{bmatrix} M_0 & M_{-1} & M_{-2} & \cdots & \cdot \\ M_1 & M_0 & M_{-1} & \cdots & \cdot \\ M_2 & M_1 & M_0 & \cdots & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & M_{-1} \\ \cdot & \cdot & \cdot & M_1 & M_0 \end{bmatrix} \times \begin{bmatrix} \cdot \\ \cdot \\ M_{-1}^{(-1)} \\ M_0^{(-1)} \\ M_1^{(-1)} \\ \cdot \end{bmatrix} = \begin{bmatrix} 0 \\ \cdot \\ 0 \\ I \\ 0 \\ \cdot \\ 0 \end{bmatrix}, \tag{66}$$

where I is the identity matrix. As observed, the block Toeplitz matrix

equation can be solved for $M_k^{(-1)}$'s, with the M_k's given in eq. (30). Having the $M_k^{(-1)}$'s and expressing eq. (36) in the form

$$[\Lambda_{-N_1}\Lambda_{-N_1+1} \cdots \Lambda_{-1}\Lambda_0\Lambda_1 \cdots \Lambda_{N_2}] \times \begin{bmatrix} M_0^{(-1)} & M_{-1}^{(-1)} & \cdots & M_{-(N_1+N_2)}^{(-1)} \\ M_1^{(-1)} & M_0^{(-1)} & \cdots & \vdots \\ \vdots & & \cdots & \vdots \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ M_{(N_1+N_2)}^{(-1)} & M_{(N_1+N_2-1)}^{(-1)} & \cdots & M_0^{(-1)} \end{bmatrix} = [-\sigma^2 M_{N_1}^{(-1)} \cdots (-\sigma^2 M_0^{(-1)} + I) \cdots -\sigma^2 M_{-N_2}^{(-1)}], \tag{67}$$

it is possible to evaluate Λ_0.

3.2 Infinite linear equalizer/decision feedback canceler

When the matrix canceler has knowledge of infinite past data symbols, it becomes a decision feedback equalizer. In addition, it may also employ a finite number of anticausal taps to operate on the future symbols, in which case it becomes a finite window canceler. This can be accomplished by a finite delay. As shown in eq. (47), to determine MSE_0, a matrix Wiener-Hopf equation has to be solved. This involves determination of anticausal factors of the $M(\theta)$ matrix as explained in Section 2.5.

There are at least two computational algorithms available for solving a matrix Wiener-Hopf equation. One method as introduced in Ref. 13 converts the matrix that has to be factored directly into a nonlinear difference equation of a Ricatti type, which converges to a stable solution. Another method, which we adopt in our present work, is a Bauer-type factorization of positive definite polynomial matrices.[14] This algorithm is suited to sampled data applications and takes advantage of the periodic and positive nature of the channel covariance matrix, $M(\theta)$, as in this work. It performs the factorization in the following steps. Suppose one desires to factor the $n \times n$ matrix $M(\theta)$ as follows:

$$M(\theta) = M^{-}(\theta)M^{+}(\theta).$$

This matrix possesses a Fourier series expansion,

$$M(\theta) = \sum_{m=-\infty}^{\infty} A_m \exp(jm\theta), \tag{68}$$

whose $n \times n$ coefficients,

$$A_m = \frac{1}{2\pi}\int_{-\pi}^{\pi} \exp(-jm\theta)M(\theta)d\theta, \tag{69}$$

approach 0_m as $|m|$ becomes infinitely large. One now follows the steps:

Step 1: Form the following variable-size Toeplitz matrix

$$T_m = \begin{bmatrix} A_0 & A_1 & \cdots & A_m \\ A_{-1} & A_0 & \cdots & A_{m-1} \\ \vdots & & & \vdots \\ A_{-m} & A_{-m+1} & \cdots & A_0 \end{bmatrix}, \tag{70}$$

of respective size $(m+1)n \times (m+1)n$, $m = 0, \cdots, \infty$ that is Hermitian and real.

Step 2: For every $m \geq 0$, perform the Cholesky's factorization[15]

$$T_m = L_m^T \cdot L_m, \tag{71}$$

where L_m^T is the transpose of L_m, and L_m is a square, real, and lower triangular matrix with positive diagonal scalar entries,

$$L_m = \begin{bmatrix} L_{00}^{(m)} & & & & & & \\ L_{10}^{(m)} & L_{11}^{(m)} & & & & & \\ \vdots & \vdots & \ddots & & & 0 & \\ & & & \ddots & & & \\ & & & & \ddots & & \\ L_{m0}^{(m)} & L_{m1}^{(m)} & \cdots & \cdots & \cdots & \cdots & L_{mm}^{(m)} \end{bmatrix}. \tag{72}$$

All blocks in eq. (72) are real $n \times n$, and all $L_{rr}^{(m)}$, $r = 0, \cdots, m$, are lower triangular.

Step 3: It has been proved in Ref. 14 that for every fixed r and k, $r \geq k \geq 0$,

$$\lim_{m \to \infty} L_{rk}^{(m)} = M_{r-k}^{+}. \tag{73}$$

The factorization of eq. (71) and the evaluation of M_{r-k}^{+} in eq. (73) are performed numerically. To find an upper bound on m for stopping the calculations, a convergence point must be established. This can be done by checking the trace of M_{r-k}^{+} in each iteration to determine whether it has reached a level of constancy and, if so, for what value of m. This completes the factorization. Indeed, in Ref. 14 it is shown that a constant trace as a function of m corresponds to a minimum of a quadratic functional:

$$I(P) = \frac{1}{2\pi} \int^{\pi} \mathrm{tr}[P^{\dagger}(\theta) M(\theta) P(\theta)] d\theta, \tag{74}$$

where $P(\theta)$ is a polynomial matrix of the form

$$P(\theta) = \sum_{r=0}^{m} \chi_r \exp(-jr\theta), \tag{75}$$

with the quadratic functional of eq. (74) expressed as

$$I(P) = \mathrm{tr}(X^{\dagger} T_m X), \tag{76}$$

where T_m was defined in eq. (70) and χ_r's represent the elements of X. Hence, there is a theoretical base for establishing the convergence point.

3.3 Finite linear equalizer/finite canceler

Finally, we consider the case where the matrix linear equalizer operates on a finite set of taps that do not overlap with those of the finite-tap matrix canceler. This is a case of great practical interest. Here, the receive filter is assumed to have a square-root-Nyquist transfer function[16] matching the transmit filter. Since it no longer matches the overall channel and transmitter characteristics, MSE_0 is a function of timing phase. Therefore, an optimum timing reference has to be established before the optimum nonstationary covariance matrix can be determined. This is accomplished here by minimizing the mean-square eye closure (MS-EC), which is a measure of the amount of received level perturbation caused by CPI and ISI.[16] In our present work it is assumed that the demodulator removes the channel phase at the optimum sampling time reference.[16] Once an optimal set of samples is found, the covariance matrix, G_{nm}, of eq. (62) is formed as

$$G_{nm} \quad n, m \in F =$$

$$\begin{bmatrix} G_{-(N_1+M_1),-(N_1+M_1)} & G_{-(N_1+M_1),-(N_1+M_1-1)} & \cdots & G_{-(N_1+M_1),(N_2+M_2)} \\ G_{-(N_1+M_1-1),-(N_1+M_1)} & & & \vdots \\ \vdots & & & \\ & & & \\ & & & \\ G_{-(N_1+1),-(N_1+M_1)} & & & \vdots \\ G_{0,-(N_1+M_1)} & & & \\ G_{(N_2+1),-(N_1+M_1)} & & & \\ \vdots & & & \vdots \\ & & & \\ G_{(N_2+M_2),-(N_1+M_1)} & \cdots\cdots\cdots\cdots & \cdots & G_{(N_2+M_2),(N_2+M_2)} \end{bmatrix} \tag{77}$$

In terms of the H_n's defined in eq. (12) the covariance matrix can be expressed as

$$G_{nm} = \sum_{l \notin S} \begin{bmatrix} H_{l+N_1+M_1} \\ \vdots \\ H_{l+N_1+1} \\ H_l \\ H_{l-N_2-1} \\ \vdots \\ H_{l-N_2-M_2} \end{bmatrix} \times [H^\dagger_{l+N_1+M_1} \cdots H^\dagger_{l+N_1+1} H^\dagger_l H^\dagger_{l-N_2-1} \cdots H^\dagger_{l-N_2-M_2}]. \tag{78}$$

Hence, by adding σ^2 to the diagonal elements of G_{nm}, the matrix R_{nm} is formed, as expressed in eq. (64). The Q_n's, that is, the coefficients of the finite window equalizer, can be computed as follows:

$$[Q_{-(N_1+M_1)} \cdots Q_{-N_1-1} Q_0 Q_{N_2+1} \cdots Q_{N_2+M_2}] = [H^\dagger_{(N_1+M_1)} H^\dagger_{(N_1+M_1-1)} \cdots H^\dagger_{-(N_2+M_2)}] \times [R_{nm}]^{-1}. \tag{79}$$

These coefficients are used in eq. (65) to determine the optimum MSE.

IV. DISCUSSION OF SIMULATIONS AND NUMERICAL RESULTS

In this section, the minimum mean-square error (MSE_0) is evaluated for the various techniques covered in the previous sections. We will first discuss a channel model, and then we will exhibit and discuss the behavior of MSE_0 as a function of the number of equalizer/canceler taps (M_1, M_2, N_1, N_2) (see Fig. 2).

4.1 Propagation model

The cross-polarization fading propagation model employed is the one that is proposed in Ref. 2 and is briefly reviewed here. The frequency characteristics of the propagation model are presented by the complex matrix

$$C(\omega) = \begin{bmatrix} C_{11}(\omega) & C_{21}(\omega) \\ C_{12}(\omega) & C_{22}(\omega) \end{bmatrix}, \tag{80}$$

where the functional form of $C_{11}(\omega)$ and $C_{22}(\omega)$ is that of a single (in line) fading channel model documented by Rummler[17] with the generic representation

$$C_{11}(\omega) = a[1 - \rho \exp(j\phi)\exp(-j\omega\tau)], \tag{81}$$

where a and ρ are real variables representing flat and dispersive fading levels, ϕ is related to the fade notch offset, and τ is the delay between

direct and reflected paths assumed to be 6.3 ns in this study. Also in the model,

$$C_{22}(\omega) = a[1 - \rho \exp(j\phi)\exp(-j(\omega - \Delta\omega)\tau)], \tag{82}$$

which is in the same form as $C_{11}(\omega)$, except for an additional variable $\Delta\omega$ that allows noncollocated fade notches to occur on the two polarization signal transfer characteristics. From Ref. 2, cross-polarized paths are assumed to behave as

$$C_{21}(\omega) = K_1 C_{11}(\omega) + K_2 C_{22}(\omega) + R_3 \exp(-j\omega D_1) \tag{83}$$

and

$$C_{12}(\omega) = K_4 C_{11}(\omega) + K_5 C_{22}(\omega) + R_6 \exp(-j\omega D_2), \tag{84}$$

where K_1, K_2, K_4, and K_5 are constants that incorporate the nonideal properties of antennas and waveguide feeds at both ends of the channel, typically taking on values varying from one hop to another in the -35 to -20 dB range. The last term in eqs. (83) and (84) represents a nondispersive cross-polarization response contributed by an independent ray. In the present work, R_3, R_6, and $\Delta\omega$ are assumed to be zero and the K_i's are assumed to be -20 dB.

4.2 Channel covariance computation

Computation of the channel covariance matrix is the initial necessary step behind all the MSE_0 calculations. In the case of the infinite window-size equalizer discussed in Sections 3.3 through 3.5, the receive filter is assumed to be a matched filter; hence, no reference timing establishment is necessary. The peak of the correlation function serves as a timing reference. By computing the sampled correlation matrix of eq. (26), we can proceed with the normalized MSE_0 calculations as explained in previous sections.

By applying the finite window equalizer, as discussed in Section 3.6, a set of optimum samples of overall impulse response is found by establishing a timing reference, t_0, for which the MS-EC of the received in-line signal is a minimum, and at this reference, the channel phase is removed.[16] This has to be done for the two polarized signals independently.

The overall transfer function matrix is given by

$$H(\omega) = C(\omega) \times P(\omega), \tag{85}$$

where $C(\omega)$ is the propagation transfer matrix and $P(\omega)$ is the diagonal Nyquist-shaping filter transfer matrix. Now, for instance, if the impulse response of the vertical in-line signal is

$$h_{i11}(t) = a[p(t) - \rho \exp(j\phi)p(t - \tau)], \tag{86}$$

where $p(t)$ is a Nyquist-shaped pulse, the channel phase becomes

$$\theta(t) = \text{Arc tg} \frac{-\rho \sin(\phi)p(t-\tau)}{p(t) - \rho \cos(\phi)p(t-\tau)}, \tag{87}$$

and the upper row block matrices of the overall impulse response matrix have to be multiplied by $\exp(-j\theta(t_0)) \times I$ (I being the unity matrix) in order to remove the channel phase at t_0. With this background, we now present the numerical results in the following subsection.

4.3 Numerical results

To provide a single set of curves for MSE_0, independent of the number of transmit states in M-QAM signal space, we normalize MSE_0 as defined in eqs. (40), (43), (46), (58a), (58b), and (65) by dividing the formulas by σ_a^2, that is, the transmitted symbols variance. In addition, we only compute the normalized MSE_0 for one of the M-QAM signals that comprise the dually polarized signals, namely, $S_v(t)$.

If one defines the unfaded signal-to-noise ratio (s/n) by Γ, it can be verified that in the case of a matched filter receiver, the normalized MSE_0 in the absence of any cross-polarization interference (K_1, K_2, K_4, K_5, R_3, $R_6 = 0$) is simply

$$\frac{1}{\sigma_a^2}\text{MSE}_0 = \frac{1}{1+\Gamma} \tag{88}$$

and, consequently, for a large unfaded s/n, it becomes Γ^{-1}. Hence, eq. (88) establishes an ultimate performance bound that can only be achieved in a utopian environment. This reference will be our baseline in the following evaluations. In a dually polarized system with a finite amount of nondispersive coupling (K_1, K_2, K_4, $K_5 > 0$), the matched filter bound is degraded somewhat. For $K_1 = K_2 = K_4 = K_5 = -20$ dB we found a small amount of degradation in the ideal MSE_0, which is not a function of the dispersive fade depth and only diminishes when there is no cross-coupling, that is, in a completely orthogonal system.

In all that follows it is assumed that the transmit filter is square-root Nyquist shaped,[16] and the receive filter either matches the overall transmitter and channel or the transmitter only. A Nyquist roll-off of 45 percent, both a 40- and a 22-MHz channel bandwidth, and an s/n of 63 dB are used in our numerical evaluations.

Figure 3 depicts the normalized MSE_0 as a function of the number of canceler taps, Q, when a 40-dB centered fade over a 22-MHz channel band is applied to both polarized signals. The linear equalizer in this case possesses an infinite number of taps. The case of pure linear equalization ($N_1 = N_2 = 0$), no cancellation, exhibits the largest MSE_0

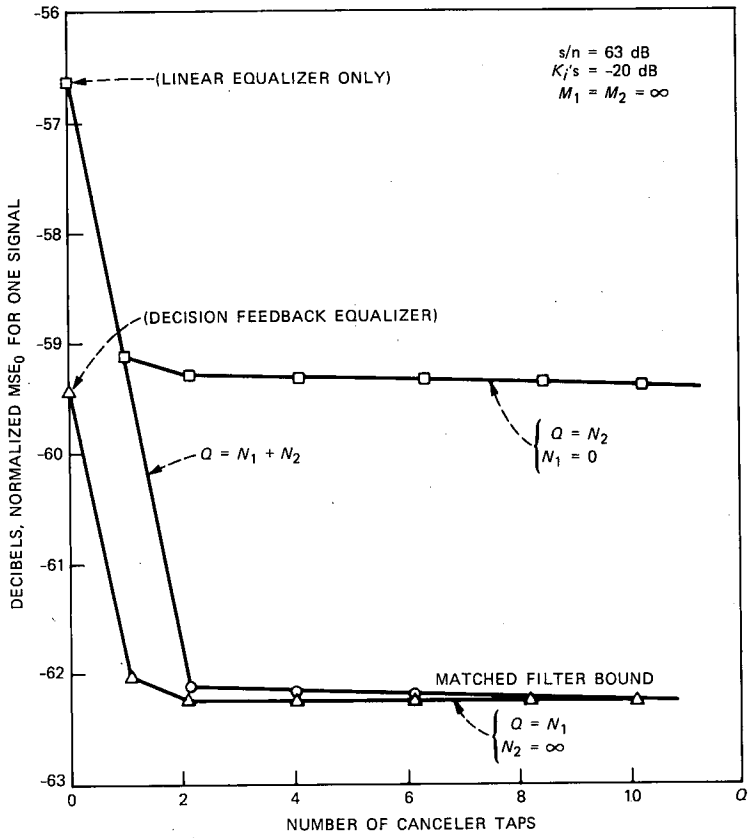

Fig. 3—Optimum normalized MSE versus number of canceler taps for a 40-dB centered fade over a 22-MHz channel.

degradation relative to the asymptotic matched filter bound. This is due to the noise enhancement experienced by the linear equalizer during deep fades. When both causal and anticausal canceler taps are present, all the curves rapidly approach the matched filter bound for a finite constant coupling ($K_i = -20$ dB, $i = 1, 2, 4, 5$). The curve for a decision feedback type canceler starts at an ideal decision feedback equalizer normalized MSE_0 and approaches the asymptotic value with two anticausal taps. The finite window size canceler curve starts at the linear equalizer case ($N_1 = N_2 = 0$) and reaches the matched filter bound asymptotic value with a total of four causal/anticausal taps. Finally, when no anticausal taps are employed the curve asymptotically approaches the ideal decision feedback case with only two causal taps.

In Fig. 4, we depict results similar to Fig. 3 for the case when the centered fade notch depth is reduced to 20 dB over a 22-MHz channel.

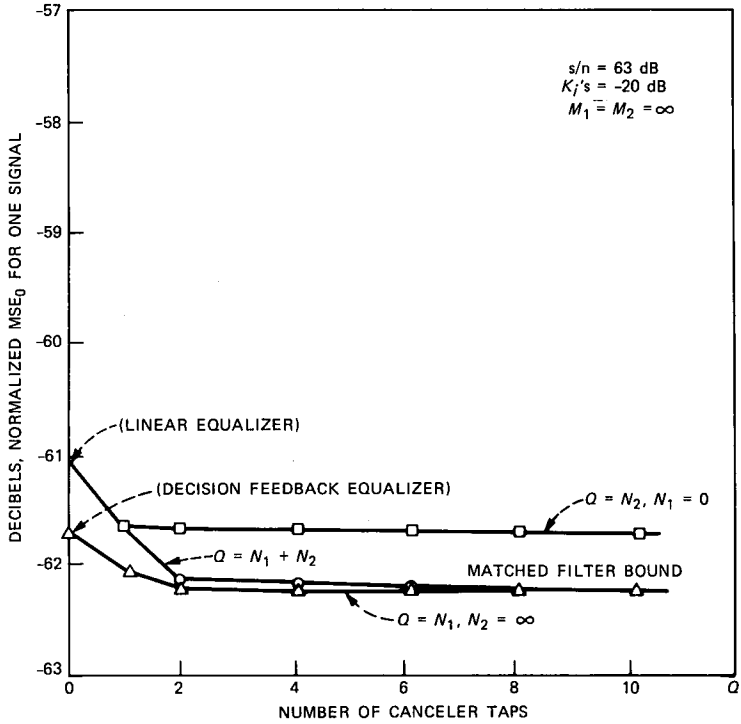

Fig. 4—Optimum normalized MSE versus number of canceler taps for a 20-dB centered fade over a 22-MHz channel.

As can be observed, the linear equalizer ($N_1 = N_2 = 0$) performance is improved. In both figures the fade notch is located at the band center; however, since in both cases the receive filter matches the overall channel and transmitter, an offset fade notch does not have a serious impact on the results for the same fade notch depth.[18]

In Figs. 5 and 6 we depict the achievable MSE_0 when the linear equalizer has a finite number of taps. The fade notch in Fig. 5 is centered, but in Fig. 6 it is offset from the band center. For ease of presenting the results in our work, fade notch offset from the band center is expressed in terms of the ratio of the fade notch distance from the band center to the channel equivalent baseband bandwidth in percentage. In Fig. 6, the fade notch is offset by 69 percent over a 22-MHz channel, that is, an offset of 7.6 MHz from the band center. As observed from Fig. 5, a total of nine taps (including the center tap) are required to achieve the asymptotic matched filter bound when decision feedback taps are present. It is interesting to note that the same asymptotic performance can be achieved no matter how the nine synchronously spaced taps are deployed between the linear equalizer

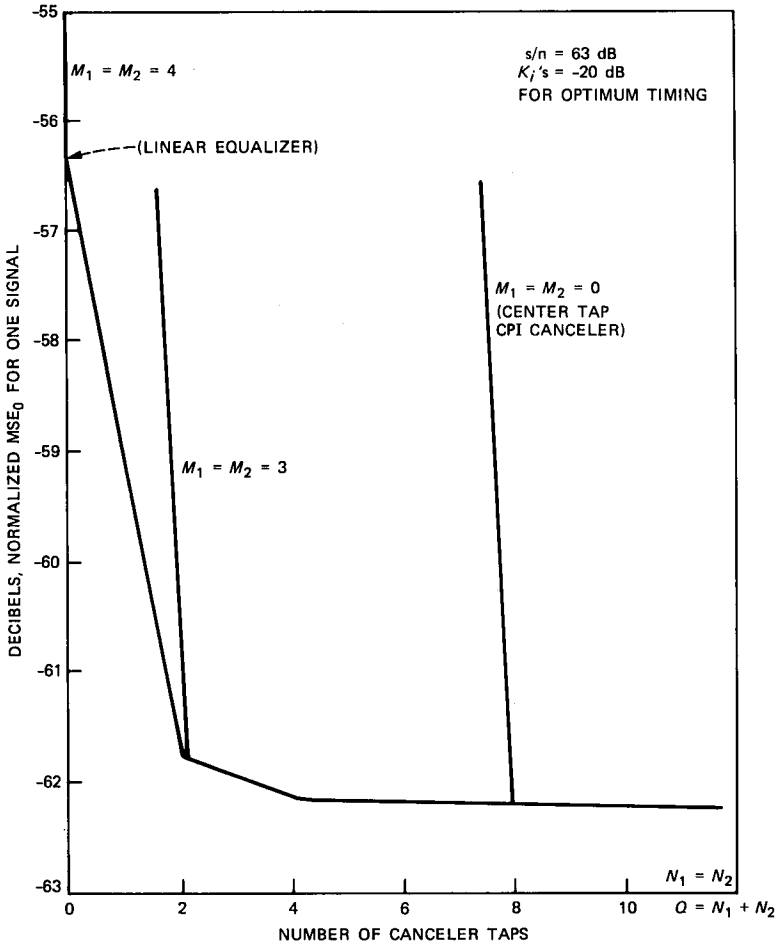

Fig. 5—Optimum normalized MSE versus number of canceler taps for a 40-dB centered fade over a 22-MHz channel.

and the canceler as long as the canceler operates in a decision feedback mode. This is because the equalizer and canceler-tap windows complement one another; therefore, since the taps do not overlap, for the same number of taps, the performance remains almost the same in the decision feedback cases. An important configuration is when the linear equalizer operates only on the main lobe of CPI by means of its center matrix taps ($M_1 = M_2 = 0$). This is a single tap linear equalizer structure as opposed to the single tap decision feedback CPI canceler proposed by Kavehrad.[3] It is clear that as long as the canceler window is sufficiently wide, a main lobe CPI canceler can achieve the asymptotic matched filter bound. The curves again indicate that deep fades degrade the linear equalizer ($N_1 = N_2 = 0$) performance significantly.

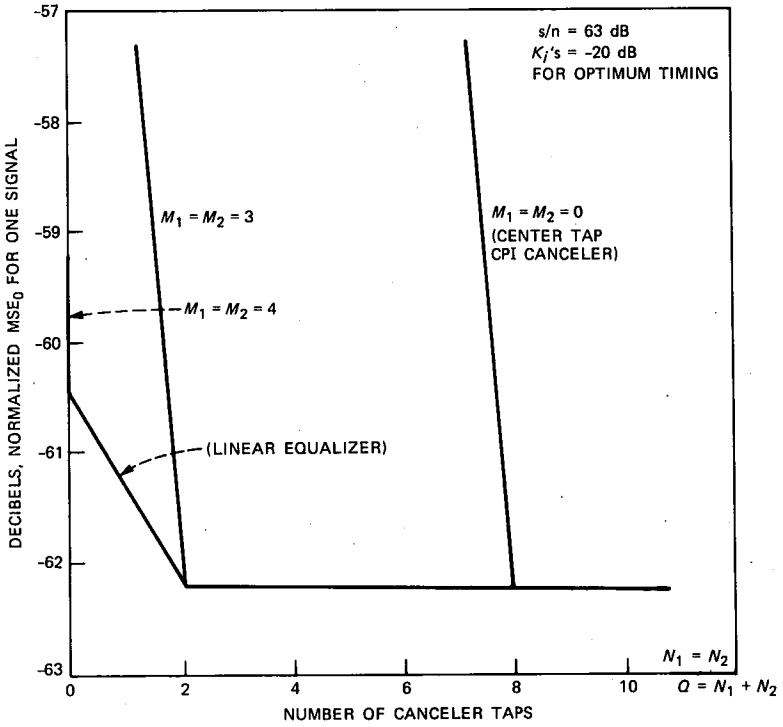

Fig. 6—Optimum normalized MSE versus number of canceler taps for a 40-dB, 69-percent offset fade over a 22-MHz channel.

Previous studies[1,2] showed that every 3-dB degradation in MSE_0 translates into a loss of 1 bit/s/Hz of data rate efficiency. Hence, linear equalization may not provide adequate rate efficiency in deep fades.

In Fig. 6 we depict curves similar to Fig. 5, except for a 40-dB offset fade with the notch frequency offset by 69 percent. Improved performance turns out to be due to the particular notch position as is brought out in the discussion of Fig. 8.

Figure 7 illustrates a similar set of curves for a 20-dB centered fading of dually polarized signals over both a 22- and a 40-MHz channel. As can be observed, the linear equalizer ($N_1 = N_2 = 0$) performance improves because of the decreased fade depth; however, over the wider channel band the degradation over decision feedback is more, as expected. This is due to the wider channel band over which the same fade notch depth causes more dispersion. The degradation amounts to 2.2-dB loss of MSE_0 comparing to a matched filter bound, that is, roughly 1 bit/s/Hz loss of data rate efficiency, and the loss can even be more for offset fades, as will be seen in Fig. 8. Hence, even

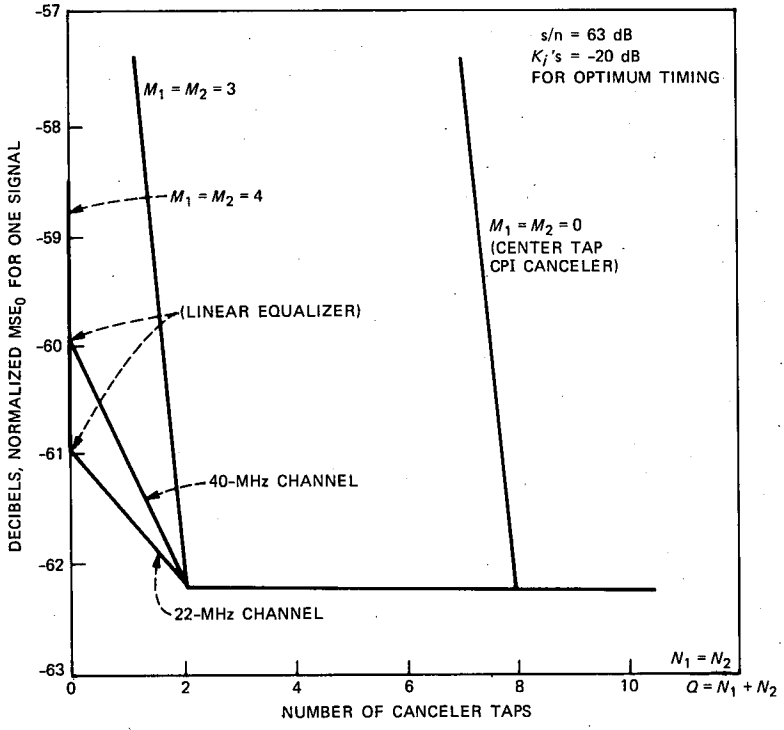

Fig. 7—Optimum normalized MSE versus number of canceler taps for a 20-dB centered fade over a 22-MHz and a 40-MHz channel.

with more typical fades the use of the linear equalizer can be troublesome over a 40-MHz channel.

Finally, to compare some of the techniques described earlier in terms of their sensitivity to fade notch offset, we plot, in Fig. 8, the normalized MSE_0 as a function of fade notch position which, as explained earlier, is expressed here in terms of the ratio of the fade notch distance from the band center to the channel equivalent baseband bandwidth. We consider the following structures:

1. A linear equalizer with

$$M_1 = M_2 = 4$$

$$N_1 = N_2 = 0 \text{ (no cancellation).}$$

2. Center tap only linear equalizer/finite window canceler with

$$M_1 = M_2 = 0$$

$$N_1 = N_2 = 4.$$

The sensitivity of the linear equalizer to offset fades is quite pro-

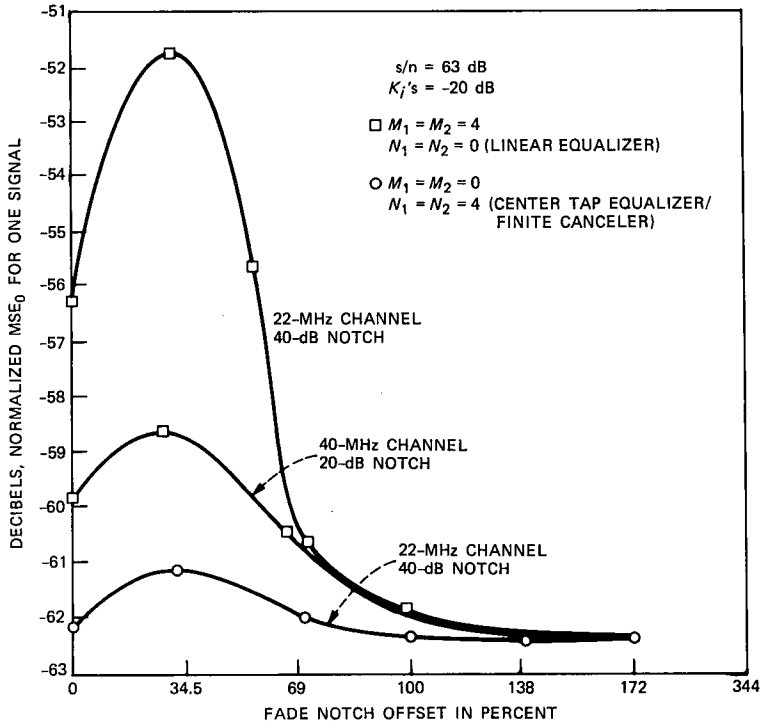

Fig. 8—Normalized MSE_0 versus fade notch offset.

nounced. The center tap equalizer with a finite window canceler exhibits a very small sensitivity to offset fades. The degradation of MSE_0 for some offset fades can be explained considering the fact that these fades cause cross-coupling of the imaginary part of a complex QAM signal into its real part, and for a particular notch offset frequency within the band, the coupling reaches its maximum. Therefore, the MSE_0 versus fade notch offset curves exhibit this phenomenon. In dually polarized systems, as in the case of the problem at hand, this is even more pronounced than in single signal transmission, because in the 4 × 4 system under offset fading there is coupling of three interfering data streams into the fourth one. A decision-feedback-type canceler structure, by canceling the major contributors to CPI and ISI and with a lesser noise enhancement, exhibits an improved performance compared with the linear equalizer. Note that all the curves in Fig. 8 have been obtained under optimum timing conditions.

To investigate the sensitivity of the two structures to timing phase, we plot in Fig. 9 the normalized MSE_0 and superimpose the normalized MS-EC of the received signal before equalization/cancellation as a function of sample timing offset from the optimum timing reference.

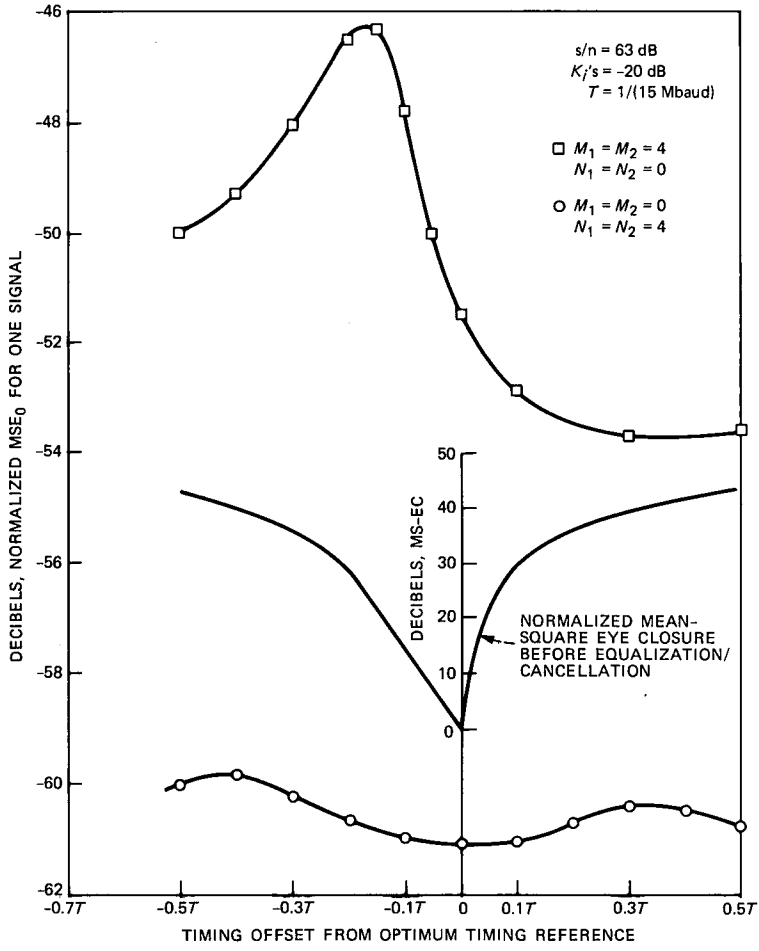

Fig. 9—Sensitivity to timing phase for a 40-dB fade notch, offset by 34.5 percent, over a 22-MHz channel.

This is done for a severe fade, namely, a 40-dB fade with a notch frequency offset by 34.5 percent over a 22-MHz channel. It is clear that the finite linear equalizer is much more sensitive to timing phase than the decision feedback type. This was previously shown in a paper by J. Salz[19] for infinite window structures in single-signal transmission. We demonstrated the concept here for dual-polarization transmission and finite window architectures. Notice that in Fig. 9 the optimum timing reference is established based on minimizing the MS-EC of the received signal in presence of fading, before CPI and ISI cancellation; hence, after cancellation occurs this timing reference may not be the one that minimizes the canceler output MSE, and

indeed the linear equalizer curve on Fig. 9 indicates this fact. As seen, the MS-EC curve has a minimum at the optimum timing reference. The sensitivity of the matrix linear equalizer to timing phase can be reduced by applying half-a-baud spaced taps, that is, by deploying fractionally spaced taps.[20] Also, a decision feedback timing method may prove more robust[21] than the minimum MS-EC timing that has been adopted in our work. Needless to say, that decision feedback timing method is more complex in terms of implementation than the minimum MS-EC timing, which can essentially be implemented at intermediate frequency. The degradation in MSE_0 seen in Fig. 9 is partly attributed to the asymmetric amplitude and delay responses of the fading channel that in the presence of a nonzero roll-off shaping filter cause a destructive addition of aliases.[16]

V. SUMMARY AND CONCLUSIONS

Current work generalizes and extends previous results[2,4] in the following respects. Data-aided decision feedback and canceler structures, known to be effective in single-channel data transmission, are adopted and included in our class of receiver structures. As a practical feature, we admit transversal filter realizations with a finite number of matrix taps and pay attention to timing phase recovery.

Because of the departure from ideal linear infinite structures considered previously, we encountered extremely difficult numerical problems, which we addressed and solved.

The dually polarized digital radio channel is modeled as a four-input port, four-output port linear network followed by additive noise. We determine the optimum admissible receiver structures when the transmitted signals are two independent M-state QAM digital data signals.

The mathematically tractable criterion, the MSE, is used throughout our work. This figure of merit has several redeeming features in addition to its being mathematically tractable. For one, it can be used to determine a sharp upper bound on error rate. More importantly, it is the quantity which is estimated in practice to provide information for updating tap coefficients in adaptive systems.

The receiver structure that minimizes the MSE consists of a matrix matched filter in cascade with a transversal filter combined with an intersymbol interference as well as a cross-polarization interference canceler. The canceler uses the detected data symbols to estimate the interference to be canceled. This is a major assumption on which our results rest. Since data-aided operations presume correct knowledge of detected data symbols and since wrong decisions will be occasionally accepted, our proposal is necessarily a boot-strapping approach. Thus, cancellation is only feasible when tentative decisions are correct most

of the time, and yet the error rate is not sufficiently low to meet system specifications. Our approach makes possible the reduction of the final error rate to an acceptable level. In circumstances where the initial error rate is very high ($\geq 10^{-2}$), our proposal will not work, and in order to ensure availability of reliable data symbols, one option is to dedicate a small fraction of the main data frame to a sequence a priori known to both transmitter and receiver for proper acquisition of data.

We use the assumption of the availability of correct data symbols to derive our main results. These are expressions for minimum attainable MSE as a function of various system parameters and numerical algorithms for evaluating the mathematical formulas—a rather intensive activity because of the large number of matrix equations that has to be solved. Inclusion of the effects of errors in the feedback/canceler loops has proved so far to be mathematically intractable.

From our extensive numerical work, which is exhibited in a sequence of graphs, we draw these major conclusions:

1. For a reasonable copolarized and cross-polarized propagation model[2] and a severe centered fade—40-dB notch depth with a secondary ray delay of 6.3 ns over an approximately 22-MHz channel bandwidth—the performance of transversal filters with a finite number of taps deployed in a decision feedback/canceler structure is substantially (6 dB) better than linear equalization, and the difference can be up to 10 dB for offset fades. It can be shown that a 3-dB increase in MSE translates into about 1 bit/s/Hz decrease in data rate at a fixed error rate or an order of magnitude increase in outage probability. Hence, linear equalization may not be adequate in deep fades. Whether this gain can be realized in practice depends on the degree of error propagation. This is difficult to assess mathematically and must be studied by computer simulation and/or by experimentation.

2. Decision feedback/canceler structures achieve the ultimate matched filter bound with only nine matrix taps provided that error propagation is neglected.

3. Nine linear equalizer taps essentially achieve the performance of the infinite-tap linear equalizer. This method is, of course, free of error propagation.

4. For milder centered fades—20-dB depth with a secondary ray delay of 6.3 ns—the linear equalizer configuration with nine taps is only 1-dB inferior to the decision feedback structure over a 22-MHz channel. However, if the channel bandwidth is increased to 40 MHz, the performance of the linear equalizer is worse than that of the decision feedback structure by 2.2 dB, and the difference can be up to 3 dB for offset fades.

5. Decision feedback/canceler configurations are less sensitive to timing phase than linear structures.

VI. ACKNOWLEDGMENTS

Discussions with N. Amitay; G. J. Foschini; L. J. Greenstein; D. J. Goodman; T. Kailath and Hanoch Lev-Ari of Stanford University; N. Kazanjian; A. M. Saleh; and D. C. Youla of Polytechnic Institute of New York were extremely valuable during the course of this work.

REFERENCES

1. G. J. Foschini and J. Salz, "Digital Communication Over Fading Radio Channels," B.S.T.J., *62,* No. 2, Part I (February 1983), pp. 429–59.
2. N. Amitay and J. Salz, "Linear Equalization Theory in Digital Data Transmission Over Dually Polarized Fading Radio Channels," AT&T Bell Lab. Tech. J., *63,* No. 10 (December 1984), pp. 2215–59.
3. M. Kavehrad, "Baseband Cross-Polarization Interference Cancellation for M-Quadrature Amplitude-Modulated Signals Over Multipath Fading Radio Channels," AT&T Tech. J., *64,* No. 8 (October 1985), pp. 1913–26.
4. M. S. Mueller and J. Salz, "A Unified Theory of Data-Aided Equalization," B.S.T.J., *60,* No. 11 (November 1981), pp. 2023–38.
5. T. Kailath, "A General Likelihood-Ratio Formula for Random Signals in Gaussian Noise," IEEE Trans. Inform. Theory, *IT-15,* No. 3 (May 1969), pp. 350–61.
6. A. Gersho and T. L. Lim, "Adaptive Cancellation of Intersymbol Interference for Data Transmission," B.S.T.J., *60,* No. 11 (November 1981), pp. 1997–2020.
7. J. Salz, "Optimum Mean-Square Decision Feedback Equalization," B.S.T.J., *52,* No. 8 (October 1973), pp. 1341–73.
8. D. C. Youla, "On the Factorization of Rational Matrices," IRE Trans. Inform. Theory, *IT-7* (July 1961), pp. 172–89.
9. N. Wiener and E. J. Akutowicz, "A Factorization of Positive Hermitian Matrices," J. Math. Mech., *8* (August 1959), pp. 111–20.
10. M. C. Davis, "Factoring the Spectral Matrix," IEEE Trans. Automat. Contr., *AC-8* (October 1963), pp. 296–305.
11. D. D. Falconer and G. J. Foschini, "Theory of Minimum Mean-Square-Error QAM Systems Employing Decision Feedback Equalization," B.S.T.J., *52,* No. 10 (December 1973), pp. 1821–49.
12. N. Levinson, "The Wiener RMS (Root Mean Square) Error Criterion in Filter Design and Prediction," J. Math. Phys., *25* (February 1946), pp. 261–78.
13. W. G. Tuel, "Computer Algorithm for Spectral Factorization of Rational Matrices," IBM J. Res. Develop., *12* (March 1968), pp. 163–70.
14. D. C. Youla and N. Kazanjian, "Bauer-Type Factorization of Positive Matrices and the Theory of Matrix Polynomials Orthogonal on the Unit Circle," IEEE Trans. Circuits Syst., *CAS-25,* No. 2 (February 1978), pp. 57–69.
15. G. H. Golub and J. H. Welsch, "Calculations of Gauss Quadrature Rules," Math. Comp., *26,* No. 106 (April 1969), pp. 221–30.
16. R. W. Lucky, J. Salz, and E. J. Weldon, *Principles of Data Communication,* New York: McGraw-Hill, 1968.
17. W. D. Rummler, "A New Selective Fading Model: Application to Propagation Data," B.S.T.J., *58,* No. 7 (May–June 1979), pp. 1037–71.
18. M. Kavehrad, "Adaptive Decision Feedback Cancellation of Intersymbol Interference Over Multipath Fading Radio Channels," Proc. ICC (June 1983), pp. 869–71.
19. J. Salz, "On Mean-Square Decision Feedback Equalization and Timing Phase," IEEE Trans. Commun., *COM-25,* No. 12, pp. 1471–76, December 1977.
20. G. Ungerboeck, "Fractional Tap-Spacing Equalizer and Consequences for Clock Recovery in Data Modems," IEEE Trans. Commun., *COM-24,* No. 8 (August 1976), pp. 856–64.
21. G. R. McMillen, M. Shafi, and D. P. Taylor, "Simultaneous Adaptive Estimation of Carrier Phase, Symbol Timing, and Data for a 49-QPRS DFE Radio Receiver," IEEE Trans. Commun., *COM-32,* No. 4 (April 1984), pp. 429–43.

AUTHORS

Mohsen Kavehrad, B.S. (Electrical Engineering), 1973, Tehran Polytechnic Institute; M.S. (Electrical Engineering), 1975, Worcester Polytechnic Institute; Ph.D. (Electrical Engineering), 1977, Polytechnic Institute of New York; Fairchild Industries, 1977–1978; GTE, 1978–1981; AT&T Bell Laboratories, 1981—. At AT&T Bell Laboratories Mr. Kavehrad is a member of the Communications Methods Research Department at Crawford Hill Laboratory. His research interests are digital communications and computer networks. He has organized and chaired sessions for IEEE sponsored conferences. Technical Editor, IEEE Communications Magazine; Chairman, IEEE Communications Chapter of New Hampshire, 1984. Member, IEEE, Sigma Xi.

Jack Salz, B.S.E.E., 1955, M.S.E., 1956, and Ph.D., 1961, University of Florida; AT&T Bell Laboratories, 1961—. Mr. Salz first worked on the electronic switching system. Since 1968 he has supervised a group engaged in theoretical studies in data communications and is currently a member of the Communications Methods Research Department. During the academic year 1967–1968, he was on leave as Professor of Electrical Engineering at the University of Florida. He was a visiting lecturer at Stanford University in Spring 1981 and a visiting MacKay Lecturer at the University of California, at Berkeley, in Spring 1983.