Suppression of Near- and Far-End Crosstalk by Linear Pre- and Post-Filtering

Michael L. Honig, Pedro Crespo, and Kenneth Steiglitz, *Fellow, IEEE*

Abstract—**Full-duplex data communications over a multi-input/multi-output linear time-invariant channel is considered. The minimum mean square error (MMSE) linear equalizer is derived in the presence of both near- and far-end crosstalk and independent additive noise, assuming correlated data, and colored noise. The MMSE equalizer is completely specified in terms of the channel and crosstalk transfer functions by using a generalization of previous work due to Salz. Conditions are given under which the equalizer can completely eliminate both near- and far-end crosstalk and intersymbol interference. The MMSE transmitter filter, subject to a transmitted power constraint, is specified when the channel and crosstalk transfer functions are bandlimited to the Nyquist frequency.**

Also considered is the design of MMSE transmitter and receiver filters when the data signals are arbitrary wide-sense stationary continuous or discrete-time signals, corresponding to the situation where the crosstalk is not phase-synchronous with the desired signal. For a particular two-input/two-output discrete-time channel model, we study the behavior of the MMSE, assuming FIR transmitter and receiver filters, as a function of how the matrix taps are allocated between these filters, and on timing phase. In this case, the jointly optimal transmitter and receiver filters are obtained numerically using an iterative technique. For the channel model considered, the MSE is a very sensitive function of timing phase, but is nearly independent of how taps are allocated between the transmitter and receiver filters.

I. Introduction

ONE of the major obstacles to realizing high data rate (i.e., at least 800 kb/s) digital subscriber lines using twisted-pair wires is crosstalk between twisted pairs in close physical proximity. Full-duplex transmission gives rise to two types of crosstalk: near-end and far-end [1]. In situations where a cable of twisted pairs is terminated at a single physical location, it has been proposed that the entire cable be treated as a single multi-input/multi-output (MIMO) channel. Crosstalk can then be characterized by the matrix impulse response of the channel, rather than as additive noise independent of the transmitted signals [2]–[4]. This work was initially concerned with the design of MIMO linear filters at the transmitter and receiver assuming full-duplex transmission, and that the transmitter and receiver on each side of the MIMO channel process the entire *vector* of inputs and outputs. In this case, crosstalk is modeled by the off-diagonal terms of the channel matrix transfer function. This work also applies, however, to the situation in which data signals on each constituent single-input/single-output (SISO) channel are phase-synchronous, but *independent* transmitters and receivers are to be designed for each SISO channel. Crosstalk in this latter situation is an interfering data signal independent of the desired signal.

The MMSE linear equalizer in the presence of synchronous interfering data signals has received attention in a number of different contexts. Perhaps the most popular application cited in the literature is multiuser communications, in which many users transmit data synchronously to various destinations over a single (SISO) channel [2], [5]–[8]. If the data signals are not orthogonal, then interference from other active users is analogous to crosstalk between twisted pairs in the subscriber loop. Other communications applications in which crosstalk is present include digital radio with diversity, in which the received signal is a vector with components from different antennas [2], [9], magnetic recording where crosstalk arises from data on adjacent tracks [10], [11], dually polarized radio channels [12], [13], and mobile and cellular radio, where crosstalk is caused by data signals from adjacent locations or cells sharing the same frequency band [14]. Related work in which crosstalk is characterized statistically as cyclostationary interference is described in [15]–[18].

The communications channel for each of the preceding applications is a special case of an N-input/M-output system, where $N, M \geq 1$. For example, assuming linear time-invariant transmission media, in the case of a single twisted pair with near-end crosstalk, the transfer function from the far-end transmitter and N synchronous near-end (interfering) transmitters to a particular receiver can be represented by a $1 \times (N + 1)$ transfer matrix, $N > 1$. Similarly, a single-input channel with diversity can be represented by an $M \times 1$ transfer matrix, $M > 1$. It appears that the structure of the MMSE equalizer for the former $1 \times (N + 1)$ case, assuming PAM data signals, was first derived in [19] and later presented in [2], [6]–[10], [17], [20], and [21] (see also [36] and the references within), although a complete specification in terms of (far-end) channel and crosstalk transfer functions is either

Manuscript received October 16, 1989; revised July 23, 1990. This work was supported in part by NSF Grant MIP-8705454 and the U.S. Army Research Office Durham Contract DAAG29-85-K-0191. This paper was presented in part at the ICC'89, Boston, MA, June 1989.

M. L. Honig is with Bellcore, Morristown, NJ 07960.

P. Crespo is with Telefonica I+D, Madrid, Spain.

K. Steiglitz is with the Department of Computer Science, Princeton University, Princeton, NJ 08544.

IEEE Log Number 9106471.

missing in these references or is not as compact as the one given here. The most general work in this area appears to be due to Salz [22] (see also [12] and [13]), who completely specified the MMSE linear equalizer for the $N \times N$ channel, assuming uncorrelated data and white noise. He also derived the MMSE transmit filter, subject to a power constraint, when the channel transfer matrix is bandlimited to the Nyquist frequency. It is pointed out in [17] that the $N \times N$ case is easily generalized to the $M \times N$ case by inserting zeroes in the appropriate entries of the channel transfer matrix. Related work is presented in [23]–[26], in which MMSE linear pre- and post-filters for discrete-time MIMO channels are derived.

In this paper, we first generalize the work of Salz [22] by deriving the MMSE MIMO linear equalizer assuming correlated data symbols, PAM data signals, and colored noise. Our derivation, given in Appendix A, differs from that given in [22] and uses classical results from sampling theory. This result is then used to specify the MMSE linear equalizer in the presence of synchronous crosstalk and additive noise in terms of near- and far-end channel transfer functions. A general condition is given in terms of these transfer functions under which crosstalk and intersymbol interference (ISI) can be completely eliminated, but at the expense of enhancing the additive noise. This "zero-forcing" condition is a special case of a zero-forcing condition for a general $N \times N$ channel, which has previously been observed in [8] and [7] (see also [17]).

The relationship between excess transmitted bandwidth and mean square error (MSE), assuming MMSE linear equalization, is also studied numerically for a specific model of a channel consisting of one twisted pair with near-end crosstalk from a single source. Our results indicate that the MMSE decreases monotonically with excess bandwidth. When crosstalk power is much greater than the power of the background noise, the use of excess bandwidth with the MMSE linear equalizer offers a significant reduction in MSE relative to the MMSE linear equalizer assuming that crosstalk is wide-sense stationary (WSS) noise.

We also consider the MMSE design of an MIMO transmitter filter, subject to a power constraint, in the presence of both near- and far-end crosstalk, ISI, and additive noise. The solution, in contrast to the corresponding MMSE equalizer, cannot be reduced to a special case of the MMSE transmitter filter for an $M \times N$ channel, and apparently has not been considered elsewhere. That is, [22], [24]–[26] consider the design of the transmitter filter in the presence of only far-end crosstalk, ISI, and additive noise. In the case of full-duplex transmission over multiple twisted pairs, however, near-end crosstalk is much more severe than far-end crosstalk. The design of transmitter and receiver FIR filters in the presence of near-end crosstalk, but using a different optimization criterion from the MSE criterion used here, is considered in [27].

Assuming the channel and crosstalk transfer functions are bandlimited to the Nyquist frequency, we obtain a necessary condition for the jointly optimal transmitter and receiver filters in the presence of near- and far-end crosstalk, ISI, and additive noise. In this case, near-end crosstalk is treated as an independent additive noise (not synchronous with the far-end transmitted signal) with spectral density that depends on the transmitted power spectral density. This problem is also equivalent to the problem of designing transmitter and receiver filters to minimize the L_2 norm between the transmitted and received signals, where the transmitted signals are assumed to be WSS with given spectra.

Finally, we consider the design of fixed-order FIR pre- and post-linear filters for a discrete-time MIMO channel in the presence of near- and far-end crosstalk, ISI, and additive noise. A two-input/two-output channel model, based upon a simplified model of two coupled twisted pairs, is used to generate numerical results for this case. A simple iterative technique is used to compute jointly optimal digital FIR transmitter and receiver filters. The MSE is found to be a sensitive function of timing phase at the receiver. The timing phases are, therefore, selected independently on each of the two channels to minimize the MSE. For the example considered, it is found that, given a fixed number of matrix taps to be allocated between the transmitter and receiver filters, the MSE is an insensitive function of how these taps are allocated (assuming that the timing phases are optimized). This is due to the assumption that near-end crosstalk energy is much greater than that of the background channel and receiver noise.

II. Optimization of Filters for PAM Data Signals

A. Channel Models

Two channel models of interest are shown in Fig. 1. In Fig. 1(a), the received signal $\boldsymbol{r}(t)$ at the bottom right consists of the sum of three components: i) the desired data signal, which is the result of passing the far-end signal $\boldsymbol{a}(t)$ through the transmitter filter $\hat{\boldsymbol{P}}$, the far-end channel $\hat{\boldsymbol{H}}$, and the receiver filter $\hat{\boldsymbol{R}}$; ii) a crosstalk signal, which is the result of passing the near-end signal $\boldsymbol{b}(t)$ through the transmitter filter, the near-end channel $\hat{\boldsymbol{G}}$, and the receiver filter; and iii) additive noise, which is $\boldsymbol{n}(t)$ filtered by $\hat{\boldsymbol{R}}$. Throughout this paper, variables underneath a "$\hat{}$" are assumed to be functions of the frequency variable f. The input vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ have $N \geq 1$ components, and the transfer functions in Fig. 1(a) are assumed to be $N \times N$ matrices. For example, this models full-duplex transmission over a cable containing $N > 1$ twisted pairs, where the transmitter (receiver) on each side of the channel processes all N channel inputs (outputs) simultaneously. In this case, $\hat{\boldsymbol{H}}$ models the far-end transmission path from one end of the cable to the other, and $\hat{\boldsymbol{G}}$ models the near-end transmission path from one end of the cable to the same side of the cable. That is, from Fig. 1(a), $[\hat{\boldsymbol{H}}]_{ij}$ is the transfer function, or frequency response, between the left side of wire i and the right side of wire j, $1 \leq i, j \leq N$ (far-end crosstalk transfer function between channel i and j), and $[\hat{\boldsymbol{G}}]_{ij}$ is the transfer function between the right side of wire i and the right side of wire j (near-

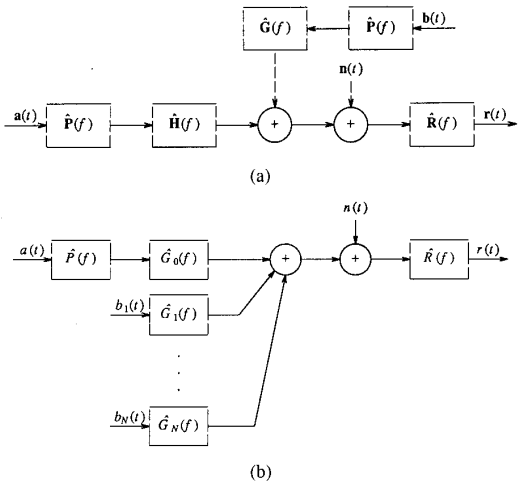


Fig. 1. (a) Simplified model for full-duplex transmission over a linear MIMO channel. (b) Channel model for $N \geq 1$ crosstalk interferers.

end crosstalk transfer function between channel i and j). Note that $[\hat{\boldsymbol{G}}]_{ii}$ is the echo transfer function on wire i.

In what follows, we will also refer to the model shown in Fig. 1(b), in which the signals and transfer functions are scalar quantities, but there may be many crosstalk signals. The data signal $a(t)$ passes through the channel modeled by $\hat{P}$, $\hat{G}_0$, and $\hat{R}$, and the ith crosstalk data signal $b_i(t)$, $1 \leq i \leq N$, passes through the near- (or far-) end crosstalk transfer function $\hat{G}_i$. This model describes the situation in which the channel is a cable containing many twisted pairs, and (scalar) transmitters and receivers are to be designed independently for each twisted pair. This situation is relevant to the subscriber loop application, since it is often quite difficult to physically coordinate transmitted signals on different twisted pairs. Of course, the two models shown in Fig. 1(a) and (b) are closely related, and it will be quite easy to modify our results for the model in Fig. 1(a) so as to apply to the model in Fig. 1(b).

In this section, we assume data signals in Fig. 1(a) of the form:

$$\boldsymbol{a}(t) = \sum_k \boldsymbol{a}_k p(t - kT) \tag{2.1a}$$

and

$$\boldsymbol{b}(t) = \sum_k \boldsymbol{b}_k p(t - kT + \theta) \tag{2.1b}$$

where $\boldsymbol{a}_k$ and $\boldsymbol{b}_k$ are N-vectors, $p(t)$ is the transmitted pulse shape, and θ is an arbitrary phase offset. The elements of $\boldsymbol{a}_k$ and $\boldsymbol{b}_k$ are complex in general, and the real and imaginary components are chosen from a discrete set of levels. Full-duplex transmission is assumed so that the receiver on the right side in Fig. 1(a) is attempting to estimate the far-end transmitted symbols $\{\boldsymbol{a}_k\}$ at the same time that the transmitter on the same side of the channel is transmitting the symbol sequence $\{\boldsymbol{b}_k\}$. The data signals in Fig. 1(b) are also assumed to be PAM signals, or scalar versions of $\boldsymbol{a}$ and $\boldsymbol{b}$ in (2.1).

B. The MMSE Linear Equalizer

Let $\boldsymbol{H}(t)$, $\boldsymbol{G}(t)$, $\boldsymbol{P}(t)$, and $\boldsymbol{R}(t)$ be the impulse responses of the far-end channel, near-end channel, transmitter filter, and receiver filter in Fig. 1(a), respectively. The output of the receiver filter $\hat{\boldsymbol{R}}$ at time kT is:

$$\boldsymbol{r}(kT) = \sum_i [\boldsymbol{R} * \tilde{\boldsymbol{H}}[(k - i)T]]\boldsymbol{a}_i + \sum_i [\boldsymbol{R} * \tilde{\boldsymbol{G}}[(k - i)T + \theta]]\boldsymbol{b}_i + \boldsymbol{R} * \boldsymbol{n}(kT) \tag{2.2}$$

where $\tilde{\boldsymbol{H}}(t) = \boldsymbol{H} * \boldsymbol{P} * p(t)$, $\tilde{\boldsymbol{G}}(t) = \boldsymbol{G} * \boldsymbol{P} * p(t)$, and "$*$" denotes convolution, i.e.,

$$\boldsymbol{P} * p(t) = \int_{-\infty}^{\infty} \boldsymbol{P}(t - s)p(s)\, ds.$$

We wish to find the receiver filter $\hat{\boldsymbol{R}}$ that minimizes the MSE, $E\{\|\boldsymbol{r}(0) - \boldsymbol{a}_0\|^2\}$. For the model in Fig. 1(b), the structure of the MMSE equalizer has been derived in [2], [6]–[10], [17], [19]–[21] assuming uncorrelated data (see also [36], which came to the authors' attention shortly before publication of this paper). It has been observed in [17] that both channel models in Fig. 1 are special cases of the model shown in Fig. 2, which shows a vector data signal $\boldsymbol{u}(t)$ as the input to an N-input/M-output channel with $M \times N$ transfer function $\hat{\boldsymbol{J}}$, where N, $M \geq 1$. That is, Fig. 2 is equivalent to Fig. 1(a) if:

$$\boldsymbol{u}(t) = \begin{bmatrix} \boldsymbol{P} * \boldsymbol{a}(t) \\ \boldsymbol{P} * \boldsymbol{b}(t) \end{bmatrix} (2N \times 1) \quad \text{and}$$

$$\hat{\boldsymbol{J}} = [\hat{\boldsymbol{H}} \quad \hat{\boldsymbol{G}}]\,(N \times 2N). \tag{2.3}$$

Furthermore, the MMSE linear equalizer for the channel in Fig. 2 has been derived by Salz [22] for the case of uncorrelated data and white noise. An abbreviated derivation of the MMSE linear equalizer for the channel model in Fig. 2 is given in Appendix A for the case of correlated data and colored noise. This equalizer is also shown in Fig. 2, and consists of the matched filter $\hat{\boldsymbol{J}}^* \boldsymbol{S}_n^{-1}$ followed by an MIMO tapped delay line with $N \times N$ matrix taps. The $M \times M$ matrix $\boldsymbol{S}_n$ is the power spectral density matrix for the additive noise, $\boldsymbol{n}(t)$, and "$*$" denotes complex conjugate transpose. (The "$\hat{\ }$" will be omitted from spectra denoted as S or $\boldsymbol{S}$.) The derivation in Appendix A is different from the one given by Salz, and uses properties of the z-transform. The transfer function of the tapped delay line is:

$$\hat{\boldsymbol{Q}}(z) = \boldsymbol{S}^d(z)[\boldsymbol{S}^J(z)\boldsymbol{S}^d(z) + \boldsymbol{I}]^{-1} \tag{2.4}$$

where for $z = e^{j2\pi f}$,

$$\boldsymbol{S}^J(e^{j2\pi f}) = \frac{1}{T}\sum_k \hat{\boldsymbol{J}}^*\left(f - \frac{k}{T}\right)\boldsymbol{S}_n^{-1}\left(f - \frac{k}{T}\right) \cdot \hat{\boldsymbol{J}}\left(f - \frac{k}{T}\right) \tag{2.5}$$

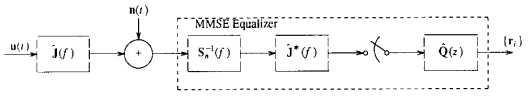


Fig. 2. MIMO channel followed by the MMSE linear equalizer.

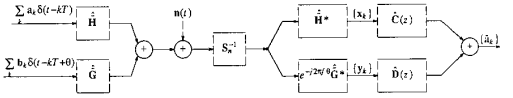


Fig. 3. MMSE receiver filter for the channel model in Fig. 1(a).

and $S^d(z)$ is the spectrum of the data sequence. That is,

$$S^d(z) = \sum_k z^{-k}\phi_k^d \tag{2.6}$$

where $\phi_k^d = E[a_i a_{i+k}^*]$. The discrete-time transfer function in (2.4) has also been recently and independently derived in [23]. Throughout this paper, a variable superscript indicates that the associated quantity refers to a discrete-time sequence (i.e., S^d and ϕ_k^d). A variable subscript is used for continuous-time processes (i.e., S_n).

We now apply the preceding result to the channel model in Fig. 1(a). When this channel model is embedded in the $M \times N$ channel $\hat{J}$ according to (2.3), the equalizer in Fig. 2 estimates both the desired data and crosstalk data. That is, the transfer function for the tapped delay line in Fig. 2 and the estimated data r_i can be partitioned as:

$$\hat{Q}(z) = \begin{bmatrix} \hat{C} & \hat{D} \\ \hat{E} & \hat{F} \end{bmatrix}, \qquad \tilde{r}_i = \begin{bmatrix} \tilde{a}_i \\ \tilde{b}_i \end{bmatrix} \tag{2.7}$$

where the $N \times N$ matrix tapped delay lines $\hat{C}$ and $\hat{D}$ operate on the sampled output of $\hat{J}^*$ to produce the estimated data sequence $\{\tilde{a}_k\}$, and $\hat{E}$ and $\hat{F}$ operate on the same sequence to produce the estimated crosstalk data sequence $\{\tilde{b}_k\}$. Notice that the receiver filter in Fig. 2, which minimizes the *total* MSE, i.e., the sum of the MSE for the desired data plus the MSE for the crosstalk data, is equivalent to the receiver that simultaneously minimizes *both* of the preceding MSE's. This is simply because the MSE for the desired (crosstalk) data is unaffected by the choice of the filters $\hat{E}$ and $\hat{F}$ ($\hat{C}$ and $\hat{D}$). Since the crosstalk data sequence is not desired, the MMSE equalizer for the channel model in Fig. 1(a) is shown in Fig. 3. That is, the MMSE equalizer has two parallel signal paths, each containing a matched filter followed by a tapped delay line. One matched filter, $\hat{H}^*$, is matched to the far-end channel transfer function, and the other, $\hat{G}^*$, is matched to the near-end channel transfer function. ($\hat{G}^*$ and $\hat{H}^*$ are the complex conjugate transpose of the Fourier transforms of $\hat{G}$ and $\hat{H}$, respectively.)

To specify $\hat{C}$ and $\hat{D}$ in terms of the channel transfer function, we first combine (2.3) and (2.5) to give:

$$S^J(z) = \begin{bmatrix} S^H(z) & S^{HG}(z) \\ S^{GH}(z) & S^G(z) \end{bmatrix} \tag{2.8}$$

where the matrix elements are aliased versions of the channel, crosstalk, and "cross-channel" spectra, i.e., for z on the unit circle,

$$S^H(e^{j2\pi f}) = \frac{1}{T}\sum_k \hat{\hat{H}}^*\left(f - \frac{k}{T}\right) S_n^{-1}\left(f - \frac{k}{T}\right)\hat{\hat{H}}\left(f - \frac{k}{T}\right) \tag{2.9a}$$

$$S^G(e^{j2\pi f}) = \frac{1}{T}\sum_k \hat{\hat{G}}\left(f - \frac{k}{T}\right) S_n^{-1}\left(f - \frac{k}{T}\right)\hat{\hat{G}}\left(f - \frac{k}{T}\right) \tag{2.9b}$$

$$S^{GH}(e^{j2\pi f}) = \frac{1}{T}\sum_k \hat{\hat{G}}^*\left(f - \frac{k}{T}\right) S_n^{-1}\left(f - \frac{k}{T}\right)\hat{\hat{H}}\left(f - \frac{k}{T}\right) \tag{2.9c}$$

and $S^{HG}(e^{j2\pi f}) = (S^{GH}(e^{j2\pi f}))^*$.

Substituting the S^J given by (2.8) into (2.4) gives:

$$\hat{C}(z) = S^d[I + S^H S^d - S^{HG}S^d(S^G S^d + I)^{-1}S^{GH}S^d]^{-1} \tag{2.10a}$$

and

$$\hat{D}(z) = S^d[S^{GH}S^d - (S^{GH}S^d + I) \cdot (S^{HG}S^d)^{-1}(S^H S^d + I)]^{-1}. \tag{2.10b}$$

Embedding the channel model in Fig. 1(b) in the model shown in Fig. 2 results in an MMSE equalizer that estimates all N interfering data sequences, in addition to the desired data sequence. Since the interfering data sequences are superfluous, the MMSE equalizer consists of $N + 1$ parallel signal paths each containing a matched filter followed by a tapped delay line, the outputs of which are summed to form an estimate of the current data symbol. The matrix S^J in this case is $(N + 1) \times (N + 1)$ with ijth element:

$$[S^J(e^{j2\pi f})]_{ij} = \frac{1}{T}\sum_k \frac{\hat{G}_{i-1}^*\left(f - \frac{k}{T}\right)\hat{G}_{j-1}\left(f - \frac{k}{T}\right)}{S_n\left(f - \frac{k}{T}\right)} \tag{2.11}$$

for $1 \le i, j \le N + 1$ and z on the unit circle. For the case of one crosstalk interferer, S^J is a 2×2 matrix, and the MMSE equalizer is shown in Fig. 3 where all transfer functions and signals are scalar quantities. From (2.10), it follows that:

$$\hat{C}(z) = \frac{(1 + S^G S^d)S^d}{(1 + S^H S^d)(1 + S^G S^d) - |S^d S^{GH}|^2} \tag{2.12a}$$

and

$$\hat{D}(z) = \frac{|S^d|^2 S^{HG}}{|S^d S^{GH}|^2 - (1 + S^G S^d)(1 + S^H S^d)} \tag{2.12b}$$

where the preceding quantities are the scalar versions of the analogous MIMO quantities, which were denoted in bold.

It is shown in Appendix A that the output MSE for the

MMSE equalizer in Fig. 2 is:

$$\text{MMSE} = \text{trace}\left\{T\int_{-1/(2T)}^{1(2T)} \boldsymbol{S}^d[\boldsymbol{S}^J\boldsymbol{S}^d + \boldsymbol{I}]^{-1}\,df\right\}. \tag{2.13}$$

For the model in Fig. 1(a), the MMSE is:

$$\text{MMSE} = \text{trace}\left\{T\int_{-1/(2T)}^{1/(2T)} \hat{\boldsymbol{C}}(e^{j2\pi f})\,df\right\} \tag{2.14}$$

where $\hat{\boldsymbol{C}}$ is given by (2.10a), or (2.12a) for the case of a single crosstalk interferer.

C. The Zero-Forcing Equalizer

Referring to the MIMO channel in Fig. 2, as the additive noise diminishes, that is, as $\sigma_n^2 \to 0$, where σ_n^2 is the variance of each noise component, then from (2.4), the transfer function of the tapped delay line $\hat{\boldsymbol{Q}}(z)$ converges to $[\boldsymbol{S}^J]^{-1}$. In the absence of additive noise, the MSE, given by (2.13), becomes zero, indicating that the equalizer can completely remove ISI and cross-channel interference (CCI), provided that $\boldsymbol{S}^J$ is nonsingular. Of course, if additive noise is present, then ISI and CCI can still be completely removed by setting $\hat{\boldsymbol{Q}} = [\boldsymbol{S}^J]^{-1}$; however, the MSE, which in this case is solely due to noise enhancement, is greater than the MMSE given by (2.13). Specifically, the zero-forcing MSE for the channel in Fig. 2 is given by

$$\text{MSE}_{\text{ZF}} = \text{trace}\left\{T\int_{-1/(2T)}^{1/(2T)} [\boldsymbol{S}^J(e^{j2\pi f})]^{-1}\,df\right\}. \tag{2.15}$$

The fact that nonorthogonal interfering data signals can sometimes be completely suppressed was first observed by Shnidman [8], who designed some specific data signals that satisfy this property. The fact that $\boldsymbol{S}^J$ must be nonsingular for the zero-forcing equalizer to exist was observed by van Etten [7]. Petersen and Falconer [17] have also observed an equivalent zero-forcing condition for the channel model in Fig. 1(b).

Referring to Fig. 1(a), in the absence of background noise the MMSE receiver can therefore perfectly reconstruct *both* the input symbol sequence $\{\boldsymbol{a}_k\}$ and the interfering symbol sequence $\{\boldsymbol{b}_k\}$ when $\boldsymbol{S}^J$ is nonsingular for all f. This is simply due to the fact that the equivalent discrete-time transfer function from the inputs $\{\boldsymbol{a}_k\}$ and $\{\boldsymbol{b}_k\}$ to the sampled outputs of the matched filters $\{\boldsymbol{x}_k\}$ and $\{\boldsymbol{y}_k\}$ in Fig. 3 is the $2N \times 2N$ matrix $\boldsymbol{S}^J(z)$. That is, in the absence of additive noise,

$$\begin{bmatrix}\boldsymbol{X}(z)\\ \boldsymbol{Y}(z)\end{bmatrix} = \begin{bmatrix}\boldsymbol{S}^H(z) & \boldsymbol{S}^{HG}(z)\\ \boldsymbol{S}^{GH}(z) & \boldsymbol{S}^G(z)\end{bmatrix}\begin{bmatrix}\boldsymbol{A}(z)\\ \boldsymbol{B}(z)\end{bmatrix} \tag{2.16}$$

where $\boldsymbol{X}(z)$, $\boldsymbol{Y}(z)$, $\boldsymbol{A}(z)$, and $\boldsymbol{B}(z)$ are the z-transforms of the sequences $\{\boldsymbol{x}_k\}$, $\{\boldsymbol{y}_k\}$, $\{\boldsymbol{a}_k\}$, and $\{\boldsymbol{b}_k\}$, respectively. Since $\boldsymbol{S}^J$ is assumed to be nonsingular for z on the unit circle, the input data sequences can be recovered from the output data sequences.

The zero-forcing transfer functions in Fig. 3 are easily obtained by computing $(\boldsymbol{S}^J)^{-1}$ from (2.8), namely

$$\hat{\boldsymbol{C}}(z) = [\boldsymbol{S}^H(z) - \boldsymbol{S}^{HG}(z)\,[\boldsymbol{S}^G(z)]^{-1}\boldsymbol{S}^{GH}(z)]^{-1} \tag{2.17a}$$

and

$$\hat{\boldsymbol{D}}(z) = [\boldsymbol{S}^{GH} - \boldsymbol{S}^G(\boldsymbol{S}^{HG})^{-1}\boldsymbol{S}^H]^{-1}. \tag{2.17b}$$

If $\hat{\tilde{\boldsymbol{G}}} = \boldsymbol{0}$ (matrix with elements equal to zero), then $\hat{\boldsymbol{C}}(z) = [\boldsymbol{S}^H(z)]^{-1}$ and $\hat{\boldsymbol{D}}(z) = 0$, which is the zero-forcing equalizer for half-duplex transmission. Also, if $\hat{\boldsymbol{G}}$ and $\hat{\boldsymbol{H}}$ are orthogonal, then $\boldsymbol{S}^{GH}(z) = \boldsymbol{S}^{HG}(z) = 0$, so that $\hat{\boldsymbol{D}}(z) = 0$ and $\hat{\boldsymbol{C}}(z) = [\boldsymbol{S}^H(z)]^{-1}$.

If $\boldsymbol{S}^J$ is singular, then in the absence of noise the MMSE linear equalizer cannot reconstruct all input data sequences to the channel. However, in the case of channels with crosstalk interference, such as those shown in Fig. 1, the receiver attempts to recover only one of the input data sequences. In this case, the MMSE receiver may be able to eliminate ISI and crosstalk even when $\boldsymbol{S}^J$ is singular. For example, referring to Fig. 1(b), this may be possible when $\hat{G}_1 = \hat{G}_2 = \cdots = \hat{G}_N$ for all f. In this case, the N crosstalk interferers appear to the receiver as a single crosstalk interferer with an expanded symbol set.

Suppose, then, that $\boldsymbol{S}^J$ corresponding to Fig. 1(b) has rank $m < N + 1$ for f in some set F, and that we wish to determine if the desired data sequence $\{a_k\}$ can be recovered in the absence of noise. For any $f_0 \in F$, we can first perform a Gaussian elimination to convert $\boldsymbol{S}^J$ to an upper triangular matrix $\boldsymbol{U}$ where the bottom $N + 1 - m$ rows contain all zeroes. Consider now the $m \times N$ submatrix of $\boldsymbol{U}$ consisting of the first m rows and the last N columns. The z-transform of the desired data sequence, $A(z)$, can be evaluated at $z = e^{j2\pi f_0}$ provided that this submatrix has rank $m - 1$. This procedure can then be used to check whether or not $A(z)$ is uniquely determined for all z on the unit circle. If so, then the desired data sequence can be recovered.

Note that the ability to eliminate ISI and CCI does not depend on the presence of matched filters in Figs. 2 and 3. Specifically, $\boldsymbol{S}_n^{-1}\hat{\boldsymbol{G}}^*$ and $\boldsymbol{S}_n^{-1}\hat{\boldsymbol{S}}^*$ can be replaced by almost any two different analog filters, $\hat{\boldsymbol{L}}$ and $\hat{\boldsymbol{M}}$, respectively, and $\hat{\boldsymbol{C}}$ and $\hat{\boldsymbol{D}}$ can, in principle, still be designed to recover $\boldsymbol{A}(z)$ in the absence of noise. Specifically, by making this substitution, (2.16) becomes:

$$\begin{bmatrix}\boldsymbol{X}(z)\\ \boldsymbol{Y}(z)\end{bmatrix} = \begin{bmatrix}\boldsymbol{S}^{LH}(z) & \boldsymbol{S}^{LG}(z)\\ \boldsymbol{S}^{MH}(z) & \boldsymbol{S}^{MG}(z)\end{bmatrix}\begin{bmatrix}\boldsymbol{A}(z)\\ \boldsymbol{B}(z)\end{bmatrix} \tag{2.18}$$

where

$$\boldsymbol{S}^{LH}(e^{j2\pi f}) = \frac{1}{T}\sum_k \hat{\boldsymbol{L}}\left(f - \frac{k}{T}\right)\hat{\boldsymbol{H}}\left(f - \frac{k}{T}\right) \tag{2.19}$$

and the remaining spectra are defined similarly. Clearly, a zero-forcing equalizer still exists provided that the matrix in (2.18) is invertible. Of course, in the presence of additive noise $\hat{\boldsymbol{L}}$ and $\hat{\boldsymbol{M}}$ should be as close as possible to $\hat{\boldsymbol{H}}^*$ and $\hat{\boldsymbol{G}}^*$, respectively, to reduce output MSE.

D. Effect of Excess Bandwidth on MSE

Consider the channel model in Fig. 1(b) for the case of a single interferer ($N = 1$). The matrix $\boldsymbol{S}^J$ is given by (2.8) and (2.9), where all quantities are scalars. It is easily verified that $\det(\boldsymbol{S}^J) = 0$ for all f if each of the sums in (2.9) contains only one nonzero term, which is the case when $\hat{G}$ and $\hat{H}$ are bandlimited to the Nyquist frequency. Consequently, in this case $\boldsymbol{S}^J$ is singular and crosstalk cannot be completely suppressed. It was originally pointed out to the authors by Falconer, Petersen, and Golden [28], [29] that if there are $N \geq 1$ interferers in Fig. 1(b), then all $N + 1$ input data sequences cannot be recovered in the absence of noise when the near- and far-end transfer functions are bandlimited to $N/(2T)$. More precisely,

$$|\hat{G}_i(f)| = 0 \qquad \text{for } f > \frac{N}{2T}, \quad i = 0, \cdots, N,$$

implies that $\boldsymbol{S}^J$ is singular *for all* f.

To prove this, note that if the components of $\hat{\boldsymbol{J}}$ are bandlimited to $N/(2T)$, then $\boldsymbol{S}^J$ is the sum of N outer products. The rank of $\boldsymbol{S}^J$ for any f is therefore less than or equal to N, which implies that $\boldsymbol{S}^J$ is singular since it has dimension $(N + 1) \times (N + 1)$. Note that, in general, the zero-forcing equalizer does not exist when $\det(\boldsymbol{S}^J) = 0$ for f in an interval of positive length.

As discussed in the preceding section, it may still be possible to recover a *single* input data sequence when $\boldsymbol{S}^J$ is singular. However, a weaker version of the preceding condition that guarantees the existence of a zero-forcing equalizer in this situation is not pursued.

For the case of one crosstalk interferer, the preceding discussion suggests that when the additive noise is small relative to crosstalk, the MSE can be significantly reduced by using excess bandwidth. The relationship between excess bandwidth and MMSE is now studied numerically for a particular model of a single twisted pair with one near-end crosstalk interferer and additive white noise. The following numerical results complement those given in [17], which assume fixed system bandwidth, but study in more detail the performance advantage obtained by treating the crosstalk as a synchronous data signal relative to treating it as a WSS random process. (See also [18], which gives numerical results for finite-complexity MMSE decision feedback equalizers.)

For the channel model in Fig. 1(b) with $N = 1$, we assume the following transfer functions, which approximately model a twisted-pair wire, and near-end crosstalk coupling from an adjacent pair [1], [30]:

$$|\hat{H}(f)|^2 = e^{-2\beta l\sqrt{f}} \tag{2.20a}$$

and

$$|\hat{G}(f)|^2 = Kf^{3/2} \tag{2.20b}$$

where l is the wire length, and β and K are constants that determine the attenuation of the far-end channel and the amount of crosstalk coupling, respectively. The following results assume $\beta = 2.3 \times 10^{-7}$ and $l = 12$ kft. The transmitter pulse is assumed to be a raised cosine, so that the combined Fourier transform of the pulse shape, transmitter filter, and front-end filter at the receiver is:

$$|\hat{P}(f)\hat{U}(f| = \begin{cases} T & 0 \leq f \leq \dfrac{1}{2T}(1-\alpha) \\ \dfrac{T}{2}\left\{1 - \sin\left[\dfrac{T\pi}{\alpha}\left(f - \dfrac{1}{2T}\right)\right]\right\} & \dfrac{1}{2T}(1-\alpha) \leq f \leq \dfrac{1}{2T}(1+\alpha) \\ 0 & f \geq \dfrac{1}{2T}(1+\alpha) \end{cases} \tag{2.21}$$

where $0 \leq \alpha \leq 1$, and α is the amount of excess bandwidth used as a fraction of the Nyquist frequency. The filter $\hat{U}$ refers to the front-end analog filter at the receiver. We assume that $\hat{U} = \hat{P}$, so that $\hat{U}$ is not matched to the channel response. The MMSE equalizer (shown in Fig. 3) follows $\hat{U}$. The equivalent noise spectrum at the input to the matched filter is, therefore, $\sigma_n^2|\hat{U}|^2$.

Fig. 4 shows plots of MSE/σ_d^2, computed from (2.14) and (2.12a), versus α for the preceding channel model, assuming uncorrelated data with variance σ_d^2, white noise, and signal-to-noise ratio $\xi = \sigma_d^2/\sigma_n^2 = 60$ dB. The baud rate $1/T = 400$ kHz, so that four-level signaling gives a rate of 800 kb/s. Also shown is MMSE versus α for the case where the crosstalk is treated as WSS additive noise with spectrum $\sigma_d^2 K f^{3/2}$. This would be the case, for instance, if the phase of the interfering data signal is random with a uniform distribution. The MMSE is then computed again from (2.14) and (2.12a), where $S^G = 0$, and $S_n = \sigma_n^2 + \sigma_d^2|\hat{G}|^2$. As will be discussed shortly, in both cases the MMSE shown in Fig. 4 can be obtained by following the front-end filter at the receiver $\hat{U}$ with an infinite-length $T/2$ fractionally-spaced equalizer. We add that the MMSE without crosstalk ($\hat{G} = 0$) is approximately -52 dB.

Two cases are shown in Fig. 4, corresponding to different values of desired signal power to crosstalk at the output of $\hat{U}$. That is, if the output of $\hat{U}$ is sampled at the rate $1/(2T)$, then the ratio of desired signal power to crosstalk power in the sampled sequence is:

$$\frac{P_d}{P_x} = \frac{\displaystyle\int_0^{(1+\alpha)/(2T)} |\hat{P}|^2|\hat{H}|^2|\hat{U}|^2\, df}{\displaystyle\int_0^{(1+\alpha)/(2T)} |\hat{P}|^2|\hat{G}|^2|\hat{U}|^2\, df}. \tag{2.22}$$

The top graph in Fig. 4 assumes $K = 10^{-10}$ in (2.20b), which gives $P_d/P_x \approx 18$ dB, and the bottom graph assumes $K = 10^{-12}$, which gives $P_d/P_x \approx 38$ dB.

Notice that the MSE corresponding to the interfering PAM signal and the MSE corresponding to WSS crosstalk are the same when $\alpha = 0$. This is because synchronous crosstalk behaves as a WSS process when sampled at the

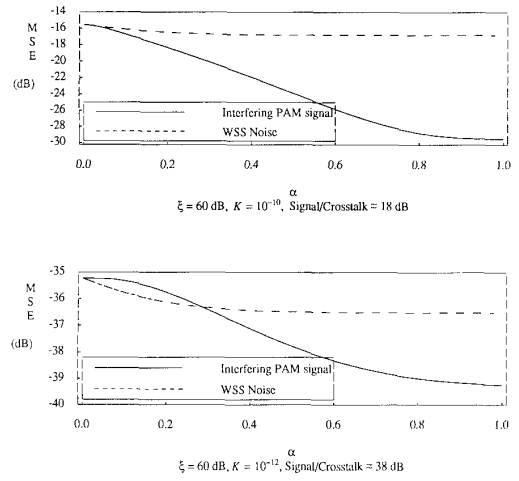


Fig. 4. Plots of MSE versus excess bandwidth α, assuming crosstalk is treated as an interfering PAM data signal, and as WSS noise.

Nyquist rate. The bottom plot shows that treating the crosstalk as a WSS process with random phase actually gives slightly less MSE than treating it as a synchronous interfering PAM signal when α is relatively small. This situation occurs only when the crosstalk power is quite small, or the ratio P_d/P_x is large.

The results in Fig. 4 indicate that the potential reduction in MSE due to the use of excess bandwidth increases with the ratio of crosstalk to background noise power. Specifically, the top plot shows a reduction in MSE in excess of 14 dB when α increases from zero to one, whereas the MSE decreases by less than 5 dB in the lower plot. This is anticipated by the preceding discussion, which indicates that if $\boldsymbol{S}^J$ is nonsingular for f in some interval, then crosstalk can be effectively suppressed in this frequency range. For the situation considered, $\boldsymbol{S}^J$ is nonsingular for $(1 - \alpha)/(2T) < |f| < 1/(2T)$. Consequently, as α increases, the wider the frequency range in which crosstalk can be suppressed. Note, however, that the zero-forcing equalizer does not exist in this example, since $\boldsymbol{S}^J$ is singular for $|f| < (1 - \alpha)/(2T)$.

E. The MMSE and Zero-Forcing Fractionally Spaced Equalizer

For a scalar (SISO) channel without crosstalk, it is well known that the MMSE linear equalizer can be implemented as a fractionally spaced equalizer (FSE) [31]. That is, the combination of a matched filter, bandlimited to $n/(2T)$, followed by a tapped-delay line can be implemented by a tapped-delay line alone, where the taps are spaced by T/n, n being a positive integer. Consider the channel and MMSE equalizer shown in Fig. 3, where all quantities are scalars. The two parallel paths in the equalizer, each consisting of a matched filter followed by a tapped-delay line, can be implemented as FSE's say FSE_G and FSE_H. The entire MMSE equalizer can therefore be replaced by a single FSE, whose tap values are the sum of the tap values of FSE_G and FSE_H. The spacing between the taps of the FSE should be no greater than T/n, assuming that both $\hat{P}\hat{G}\hat{U}$ and $\hat{P}\hat{H}\hat{U}$ are bandlimited to $n/(2T)$.

The preceding discussion implies that the MMSE equalizer for the channel in Fig. 1(b) can be implemented as a single FSE, where the spacing between taps is no greater than T/n, n being the smallest integer such that $\hat{P}\hat{G}_i\hat{U} = 0$, $|f| > n/(2T)$, $i = 0, \cdots, N$. From the discussion in the last section, if a zero-forcing equalizer exists for the channel in Fig. 1(b), then the bandwidth of $\hat{P}\hat{G}_i\hat{U}$ must be at least $(N + 1)/(2T)$ for some i. Consequently, the zero-forcing equalizer can be implemented as an infinite-length FSE where the spacing between taps is no greater than $T/(N + 1)$. Of course, in practice the FSE must have a finite number of taps, which are adapted to minimize the MSE.

F. MMSE Transmitter Filter

In general, a closed-form solution for the transmitter filter, $\hat{\boldsymbol{P}}$, that minimizes the MSE in Fig. 1(a) or (b), subject to a power constraint and assuming the MMSE receiver filter, seems quite difficult to obtain. It is possible to derive a necessary condition for $\hat{P}$ in Fig. 1(b) by applying a variational argument to the expression for MMSE, (2.14). However, the resulting expression is complicated and does not seem to yield much insight.

The discussion in Section II-D indicates that improved MMSE performance can be obtained by using transmitter filters with bandwidths in excess of the Nyquist interval $[-1/(2T), 1/(2T)]$, which is in contrast to the case when crosstalk is not present [32, Sect. 7.3.3]. Nevertheless, if we assume that the transfer function $\hat{\boldsymbol{J}}$ is bandlimited to the Nyquist interval, then at the sampling instants the crosstalk becomes a WSS additive noise. The total noise spectrum is then:

$$S_{\text{tot}}(f) = S_n(f) + S^d(e^{j2\pi f}) \sum_{i=1}^{N} |\hat{G}_i(f)|^2. \quad (2.23)$$

The MMSE receiver consists of the matched filter $\hat{G}_0/S_{\text{tot}}$ followed by a tapped-delay line. The MMSE transmitter filter can be obtained by applying a variational argument to (2.14), where the power constraint

$$\frac{1}{T}\int_{-1/(2T)}^{1/(2T)} S^d(e^{j2\pi f})|\hat{P}(f)|^2\, df = P_{\text{trans}} \quad (2.24)$$

is taken into account via a Lagrange multiplier μ. The result is:

$$|\hat{P}|^2 = \frac{S_{\text{tot}}}{S^d}\,\frac{\sqrt{\dfrac{S^d}{\mu S_{\text{tot}}}}\,|\hat{G}_0| - 1}{|\hat{G}_0|^2} \qquad \text{for } f \in F \quad (2.25a)$$

and $|\hat{P}| = 0$ for $f \notin F$, where:

$$F(f) = \left\{ f\colon \sqrt{\frac{S^d}{\mu S_{\text{tot}}}}\,|\hat{G}_0| \geq 1 \right\}. \quad (2.25b)$$

This expression reduces to the MMSE transmitter filter given in [32, Sect. 7.3.3] when the data is uncorrelated, i.e., $S^d(e^{j2\pi f}) = \sigma_d^2$, and $\hat{G}_i = 0$, $i = 1, \cdots, N$ ($S_{\text{tot}} = S_n$).

Optimization of the transmitter filter in Fig. 1(a) is deferred until the next section on continuous waveform estimation. We point out, however, that optimizing the transmitter filter in Fig. 1(a) gives a different solution from the one just obtained for Fig. 1(b), since the *same* transmitter filter precedes *both* the far-end and near-end transfer functions. This would be the case if an MIMO transmitter filter is attached to an MIMO full-duplex channel. If all signals and transfer functions in Fig. 1(b) are scalars, then the situation modeled is full-duplex transmission over a SISO channel with echo.

In addition to the transmitter optimization problems considered here, one could also extend the model in Fig. 1(b) by assuming all of the crosstalk transfer functions are preceded by different transmitter filters. Taking into account that these transmitter filters are also connected to far-end channels as well, the problem is then to minimize an appropriate cost function (for example, total MSE over all channels) subject to transmitted power constraints. Such an optimization problem appears to be quite complicated, and is not pursued further.

III. Continuous-Time Problem

In this section, we optimize the transmitter and receiver filters for the case where the data signals are arbitrary WSS waveforms. The filters are designed to minimize the L_2 norm between the transmitted and received signals. This situation applies to analog communications systems with crosstalk, and also data communications systems in which the interfering data signals are not synchronous with the received signal.

From Fig. 1(a), the received signal vector is:

$$\boldsymbol{r}(t) = \boldsymbol{R} * \boldsymbol{H} * \boldsymbol{P} * \boldsymbol{a}(t) + \boldsymbol{R} * \boldsymbol{G} * \boldsymbol{P} * \boldsymbol{b}(t) + \boldsymbol{R} * \boldsymbol{n}(t) \tag{3.1}$$

where $\boldsymbol{P}(t)$, $\boldsymbol{R}(t)$, $\boldsymbol{H}(t)$, and $\boldsymbol{G}(t)$ are the matrix impulse responses of the transmitter and receiver filters, and far- and near-end channels, respectively. Taking the Fourier transform of both sides of (3.1) gives

$$\hat{\boldsymbol{r}}(f) = \hat{\boldsymbol{R}}(f)\hat{\boldsymbol{H}}(f)\hat{\boldsymbol{P}}(f)\hat{\boldsymbol{a}}(f) + \hat{\boldsymbol{R}}(f)\hat{\boldsymbol{G}}(f)\hat{\boldsymbol{P}}(f)\hat{\boldsymbol{b}}(f) + \hat{\boldsymbol{R}}(f)\hat{\boldsymbol{n}}(f) \tag{3.2}$$

where the Fourier transform of the vector function $\boldsymbol{V}(t)$ is denoted as $\boldsymbol{V}(f)$.

The MSE can be written as:

$$E\left(\int_{-\infty}^{\infty} \|\hat{\boldsymbol{r}} - \hat{\boldsymbol{a}}\|^2 \, df\right) = E\left(\int_{-\infty}^{\infty} \|[\hat{\boldsymbol{R}}\hat{\boldsymbol{H}}\hat{\boldsymbol{P}} - \boldsymbol{I}]\boldsymbol{a} + \hat{\boldsymbol{R}}\hat{\boldsymbol{G}}\hat{\boldsymbol{P}}\hat{\boldsymbol{b}} + \hat{\boldsymbol{R}}\hat{\boldsymbol{n}}\|^2 \, df\right) \tag{3.3}$$

where the expectation is taken with respect to $\boldsymbol{b}(t)$ and $\boldsymbol{a}(t)$. Assuming that $\boldsymbol{b}$ and $\boldsymbol{a}$ are independent, but have the same second-order statistics, the expectation in (3.3) can be evaluated as:

$$\text{MSE}\,(\hat{\boldsymbol{R}}, \hat{\boldsymbol{P}}) = \text{trace}\left\{\int_{-\infty}^{\infty} [(\hat{\boldsymbol{R}}\hat{\boldsymbol{H}}\hat{\boldsymbol{P}} - \boldsymbol{I})\boldsymbol{S}_d(\hat{\boldsymbol{R}}\hat{\boldsymbol{H}}\hat{\boldsymbol{P}} - \boldsymbol{I})^* + \hat{\boldsymbol{R}}\hat{\boldsymbol{G}}\hat{\boldsymbol{P}}\boldsymbol{S}_d(\hat{\boldsymbol{R}}\hat{\boldsymbol{G}}\hat{\boldsymbol{P}})^* + \hat{\boldsymbol{R}}\boldsymbol{S}_n\hat{\boldsymbol{R}}^*] \, df\right\} \tag{3.4}$$

where the $N \times N$ matrices $\boldsymbol{S}_d$ and $\boldsymbol{S}_n$ are the spectra of the data (desired) and noise (N-vector) waveforms, respectively. Note that "n" and "d" are subscripts, so as to distinguish these spectra from the data and noise spectra in Section II. For fixed $\hat{\boldsymbol{P}}$, the MMSE receiver filter has been derived in [3], assuming the desired signal and noise are white. If the designed signal and noise are colored, the MMSE receiver filter is:

$$\hat{\boldsymbol{R}}_{\text{opt}} = \boldsymbol{S}_d\hat{\boldsymbol{P}}^*\hat{\boldsymbol{H}}^*\hat{\boldsymbol{A}}^{-1} \tag{3.5}$$

where

$$\hat{\boldsymbol{A}} = \hat{\boldsymbol{H}}\hat{\boldsymbol{P}}\boldsymbol{S}_d\hat{\boldsymbol{P}}^*\hat{\boldsymbol{H}}^* + \hat{\boldsymbol{G}}\hat{\boldsymbol{P}}\boldsymbol{S}_d\hat{\boldsymbol{P}}^*\hat{\boldsymbol{G}}^* + \boldsymbol{S}_n. \tag{3.6}$$

Note that this receiver filter is simply the standard Wiener filter for a channel with transfer function $\hat{\boldsymbol{H}}$ and additive WSS noise with spectrum $\boldsymbol{S}_n + \hat{\boldsymbol{G}}\hat{\boldsymbol{P}}\boldsymbol{S}_d\hat{\boldsymbol{P}}^*\hat{\boldsymbol{G}}^*$. Substituting for $\hat{\boldsymbol{R}}$ in (3.4) gives:

$$\text{MSE}\,(\hat{\boldsymbol{R}}_{\text{opt}}, \hat{\boldsymbol{P}}) = \text{trace}\int_{-\infty}^{\infty} [\boldsymbol{S}_d(\boldsymbol{I} - \hat{\boldsymbol{P}}^*\hat{\boldsymbol{H}}^*\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{H}}\hat{\boldsymbol{P}}\boldsymbol{S}_d)] \, df. \tag{3.7}$$

Assuming for the moment that the desired signal and noise are white, i.e., $\boldsymbol{S}_d = \sigma_d^2\boldsymbol{I}$ and $\boldsymbol{S}_n = \sigma_n^2\boldsymbol{I}$, then (3.7) can be rewritten as:

$$\text{MSE}\,(\hat{\boldsymbol{R}}_{\text{opt}}, \hat{\boldsymbol{P}}) = \sigma_d^2\,\text{trace}\left\{\int_{-\infty}^{\infty}\left[\boldsymbol{I} + \hat{\boldsymbol{H}}\hat{\boldsymbol{W}}\hat{\boldsymbol{H}}^* \cdot \left(\frac{1}{\xi}\boldsymbol{I} + \hat{\boldsymbol{G}}\hat{\boldsymbol{W}}\hat{\boldsymbol{G}}^*\right)^{-1}\right]^{-1} df\right\} \tag{3.8}$$

where $\hat{\boldsymbol{W}} = \hat{\boldsymbol{P}}\hat{\boldsymbol{P}}^*$, and the signal-to-noise ratio $\xi = \sigma_d^2/\sigma_n^2$.

It is easily verified that the expression for the MMSE receiver filter, (3.5), is the same as the MMSE receiver filter for a PAM data signal derived in Section II when the near- and far-end channels in the latter case are bandlimited to the Nyquist frequency $1/(2T)$. The expression for MSE, (3.7), is then the same as the analogous expression for MSE in the PAM case. If the near-end impulse response $\boldsymbol{G}(t) = 0$, then the expression for $\hat{\boldsymbol{R}}_{\text{opt}}$ and the associated MSE reduce to results previously derived in [24]–[26].

We now wish to minimize the expression for MSE given by (3.7) over $\hat{\boldsymbol{P}}$ subject to an average transmitted power constraint, i.e.,

$$\text{trace}\left\{\int_{-\infty}^{\infty} \hat{\boldsymbol{P}}\boldsymbol{S}_d\hat{\boldsymbol{P}}^* \, df\right\} = P_{\text{trans}}. \tag{3.9}$$

A necessary condition for the optimal transmitter transfer matrix can be derived by applying a variational argument to (3.7). A simpler approach, however, is to first fix the receiver filter, and solve for the $\hat{\boldsymbol{P}}$ that minimizes the MSE subject to the power constraint (3.9). The solution is:

$$\hat{\boldsymbol{P}}_{\text{opt}} = (\hat{\boldsymbol{H}}^*\hat{\boldsymbol{R}}^*\hat{\boldsymbol{R}}\hat{\boldsymbol{H}} + \hat{\boldsymbol{G}}^*\hat{\boldsymbol{R}}^*\hat{\boldsymbol{R}}\hat{\boldsymbol{G}} + \mu\boldsymbol{I})^{-1}\hat{\boldsymbol{H}}^*\hat{\boldsymbol{R}}^*. \quad (3.10)$$

It is shown in Appendix B that combining (3.10) with (3.5) gives the following necessary condition for the optimal transmitter transfer characteristic, $\hat{\boldsymbol{W}} = \hat{\boldsymbol{P}}\hat{\boldsymbol{P}}^*$, when the data signal is white, i.e., $\boldsymbol{S}_d = \sigma_d^2\boldsymbol{I}$,

$$\sigma_d^2\hat{\boldsymbol{H}}^*\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{G}}\hat{\boldsymbol{W}}\hat{\boldsymbol{G}}^*\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{H}} - \sigma_d^2\hat{\boldsymbol{G}}^*\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{H}}\hat{\boldsymbol{W}}\hat{\boldsymbol{H}}^*\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{G}}$$

$$+ \hat{\boldsymbol{H}}^*\hat{\boldsymbol{A}}^{-1}\boldsymbol{S}_n\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{H}} = \frac{\mu}{\sigma_d^2}\boldsymbol{I} \quad (3.11)$$

where μ is the Lagrange multiplier chosen to satisfy the constraint (3.9). Since $\hat{\boldsymbol{W}}$ is real and symmetric, it can be written as $\hat{\boldsymbol{W}} = \Phi(f)\boldsymbol{V}(f)\Phi^*(f)$ where Φ is an orthonormal matrix for each f, and $\boldsymbol{V}$ is diagonal. Consequently, the optimal transmitter filter has the form

$$\hat{\boldsymbol{P}} = \Phi\boldsymbol{V}^{1/2}\boldsymbol{U} \quad (3.12)$$

where $\boldsymbol{U}$ is an orthonormal matrix. Note that $\boldsymbol{U}$ can be selected to allocate power among the constituent SISO channels. Algorithms that compute the $\boldsymbol{U}$ to satisfy specified power constraints on each constituent channel are discussed in [12] for the two-input/two-output case, and in [24] and [25] for the general N-input/N-output case.

In the absence of echo and near-end crosstalk and assuming that $\boldsymbol{S}_n = \sigma_n^2\boldsymbol{I}$, the optimality condition (3.11) simplifies to:

$$\frac{1}{\xi}\hat{\boldsymbol{H}}^*(\hat{\boldsymbol{H}}\hat{\boldsymbol{W}}\hat{\boldsymbol{H}}^* + \frac{1}{\xi}\boldsymbol{I})^{-2}\hat{\boldsymbol{H}} = \mu\boldsymbol{I}. \quad (3.13)$$

Assuming that $\hat{\boldsymbol{H}}\hat{\boldsymbol{H}}^*$ is nonsingular, this can be written as:

$$\frac{1}{\xi\mu}\hat{\boldsymbol{H}}\hat{\boldsymbol{H}}^* = \left(\hat{\boldsymbol{H}}\hat{\boldsymbol{W}}\hat{\boldsymbol{H}}^* + \frac{1}{\xi}\boldsymbol{I}\right)^2. \quad (3.14)$$

It is shown in Appendix B that a solution to (3.14) is given by:

$$\hat{\boldsymbol{W}} = \frac{1}{\sqrt{\mu\xi}}(\hat{\boldsymbol{H}}^*\hat{\boldsymbol{H}})^{-1/2} - \frac{1}{\xi}(\hat{\boldsymbol{H}}^*\hat{\boldsymbol{H}})^{-1}. \quad (3.15)$$

According to the preceding discussion, this expression for the MMSE transfer function should be the same as the MMSE transmitter transfer function for PAM data transmission when the channel is bandlimited to the Nyquist frequency. That is, (3.15) gives the MMSE transmitter transfer characteristic for half-duplex PAM transmission over an MIMO channel, which was first obtained by Amitay and Salz [12], [22]. It is shown in Appendix B that this expression is indeed the same as the analogous result in [22].

The only case for which we can derive a closed-form expression for $\hat{\boldsymbol{W}}$ is when both $\hat{\boldsymbol{G}}$ and $\hat{\boldsymbol{H}}$ are diagonal. This corresponds to the presence of only echo and ISI. Assuming arbitrary data and noise spectra, the MMSE transmitter filter for a SISO full-duplex channel is given by:

$$|\hat{\boldsymbol{P}}|^2 = \frac{S_n}{S_d}\frac{\sqrt{\frac{S_d}{S_n\mu}}|\hat{H}| - 1}{|\hat{G}|^2 + |\hat{H}|^2} \qquad \text{for } f \in F \quad (3.16a)$$

and $|\hat{P}|^2 = 0$ when $f \notin F$, where $\hat{H}$ and $\hat{G}$ are the SISO far-end and near-end (echo) channel transfer functions, respectively, and:

$$F = \left\{f: \sqrt{\frac{S_d}{\mu S_n}}|\hat{H}| \geq 1\right\}. \quad (3.16b)$$

Note that (3.16) reduces to the MMSE transmitter filter given by (2.25) when $\hat{G} = 0$. Computing the Lagrange multiplier μ from (3.9) gives:

$$\mu = \left[\frac{\int_F \frac{\sqrt{S_dS_n}|\hat{H}|}{|\hat{G}|^2 + |\hat{H}|^2}\,df}{P_{\text{trans}} + \int_F \frac{S_n}{|\hat{G}|^2 + |\hat{H}|^2}\,df}\right]^2. \quad (3.17)$$

Since F depends on μ, in general μ must be obtained by an iterative numerical technique. Substituting (3.16) into (3.5) relates the MMSE receiver filter to the MMSE transmitter filter:

$$\hat{R}_{\text{opt}} = \sqrt{\frac{S_d\mu}{S_n}}\frac{\hat{P}^*\hat{H}^*}{|\hat{H}|}, \qquad f \in F \quad (3.18a)$$

or

$$|\hat{R}|^2 = \frac{S_d\mu}{S_n}|\hat{P}|^2, \qquad f \in F \quad (3.18b)$$

which is independent of $\hat{G}$. Finally, substituting for $|\hat{P}|^2$ in (3.7) gives the minimum SISO MSE when both the transmitter and receiver filters are optimized:

$$\text{MSE}_{\min} = \int_F S_d\frac{|\hat{G}|^2 + \sqrt{\frac{\mu S_n}{S_d}}|\hat{H}|}{|\hat{G}|^2 + |\hat{H}|^2}\,df + \int_{f\notin F} S_d\,df. \quad (3.19)$$

If the noise is white, i.e., $S_n = \sigma_n^2$, then (3.16) and (3.17) can be combined to give:

$$|\hat{P}|^2 = \left(\int_F \frac{\sqrt{S_d}|\hat{H}|}{|\hat{G}|^2 + |\hat{H}|^2}\,df\right)^{-1}\frac{P_{\text{trans}}|\hat{H}|}{\sqrt{S_d}(|\hat{G}|^2 + |\hat{H}|^2)}$$

$$+ \frac{\sigma_n^2}{\sqrt{S_d}(|\hat{G}|^2 + |\hat{H}|^2)}$$

$$\cdot\left[|\hat{H}|\frac{\int_F \frac{1}{|\hat{G}|^2 + |\hat{H}|^2}\,df}{\int_F \frac{\sqrt{S_d}|\hat{H}|}{|\hat{G}|^2 + |\hat{H}|^2}\,df} - \frac{1}{\sqrt{S_d}}\right]. \quad (3.20)$$

As $\sigma_n^2 \to 0$, leaving ISI and echo as the only channel impairments, the last term on the right disappears, and the set F tends to $(-\infty, \infty)$. Conversely, as the background noise level represented by σ_n^2 increases, the second term on the right dominates. Letting $\sigma_n^2 \to 0$, and combining this with (3.17) and (3.18) gives the magnitude of the optimal receiver transfer function:

$$|\hat{R}|^2 = \left(\int_{-\infty}^{\infty} \frac{\sqrt{S_d}|\hat{H}|}{|\hat{G}|^2 + |\hat{H}|^2}\, df \right)^{-1} \frac{\sqrt{S_d}|\hat{H}|}{P_{\text{trans}}(|\hat{G}|^2 + |\hat{H}|^2)}. \tag{3.21}$$

Note that $|\hat{P}|\,|\hat{R}|$ in this case is independent of the transmitter power P_{trans}. This is due to the fact that increasing P_{trans} also increases ISI and echo, so that the effective signal-to-noise ratio is independent of the transmitted power. The minimum MSE in this case is:

$$\text{MSE}_{\min} = \int_{-\infty}^{\infty} \frac{S_d|\hat{G}|^2}{|\hat{G}|^2 + |\hat{H}|^2}\, df \tag{3.22}$$

which is also independent of the transmitted power. If, however, $\sigma_n^2 > 0$, then it can be shown that the MSE is a decreasing function of P_{trans}. Note that by setting $\hat{G} = 0$, the previous expressions reduce to the corresponding known results for half-duplex transmission.

IV. Discrete-Time Problem

Consider again the transmission of PAM data signals in the presence of crosstalk, as shown in Fig. 1(a). Suppose that the receiver filter consists of an analog front-end filter followed by a T-spaced tapped-delay line. This may be the case in situations where the MMSE filter is too expensive to implement. The channel then has an equivalent discrete-time description [30] in which the crosstalk can be treated as WSS. We can again consider the problem of designing MMSE discrete-time receiver and transmitter filters for this equivalent discrete-time channel in the presence of crosstalk. If the transmitter and receiver filters are assumed to be IIR (infinite complexity), then the problem is completely isomorphic to that considered in Section III and the same results follow without change, except that: Fourier transforms are replaced by z-transforms with $z = e^{j2\pi f}$; integrals on the imaginary axis are replaced by integrals on the unit circle.

A. The Discrete-Time Fixed-Order FIR Case

To establish the notation, we begin with the simple case of optimizing the receiver with fixed transmitter filter and no crosstalk. The transmitter filter will be represented by a sequence $\{\boldsymbol{P}_k, 0 \le k \le m\}$ of $N \times N$ matrices, where the elements of the kth matrix represent the impulse response of the filter at time k. Similarly, the channel filter is represented by a sequence of matrices $\{\boldsymbol{H}_k, 0 \le k \le m\}$ and the receiver filter by a sequence $\{\boldsymbol{R}_k, 0 \le k \le l\}$.

To write the output as a matrix operating on the input, we define the following matrix constructions. Let $\{\boldsymbol{A}_k, 0 \le k \le p\}$ be a sequence of $N \times N$ matrices, $N \ge 1$. Then the *row-expanded* matrix $\boldsymbol{A}\_$ is:

$$\boldsymbol{A}\_ = [\boldsymbol{A}_0|\boldsymbol{A}_1|\boldsymbol{A}_2|\cdots|\boldsymbol{A}_p]. \tag{4.1}$$

The sequence index k of $\{\boldsymbol{A}_k\}$ is regarded as a time parameter, and we can view this matrix as a time expansion. It has N rows and $(p+1)N$ columns. We always assume that any postmultiplier is compatible in its dimensions with $\boldsymbol{A}\_$. Similarly, the *column-expansion* of the sequence $\{\boldsymbol{B}_k, 0 \le k \le q\}$ of $N \times N$ matrices is defined by:

$$\boldsymbol{B}_| = \begin{bmatrix} \boldsymbol{B}_q \\ \boldsymbol{B}_{q-1} \\ \boldsymbol{B}_{q-2} \\ \cdots \\ \boldsymbol{B}_0 \end{bmatrix} \tag{4.2}$$

which has $(q+1)N$ rows and N columns. Again, this is regarded as a time expansion, where q is the current time index, and any premultiplier is assumed compatible with the dimensions of $\boldsymbol{B}_|$. Note that time is shifted backward, so that matrix-vector products correspond to convolutions. Specifically, the convolution of the sequence $\{\boldsymbol{A}_k\}$ with $\{\boldsymbol{B}_k\}$ can be expressed as $\boldsymbol{A}\_[\boldsymbol{B}\_]_|$ where $[\boldsymbol{B}\_]_|$ is an $N(p+1) \times N(q+p+1)$ convolution matrix. For example, if $p = 3$ and $q = 2$,

$$[\boldsymbol{B}\_]_| = \begin{bmatrix} \boldsymbol{B}_0 & \boldsymbol{B}_1 & \boldsymbol{B}_2 & 0 & 0 & 0 \\ 0 & \boldsymbol{B}_0 & \boldsymbol{B}_1 & \boldsymbol{B}_2 & 0 & 0 \\ 0 & 0 & \boldsymbol{B}_0 & \boldsymbol{B}_1 & \boldsymbol{B}_2 & 0 \\ 0 & 0 & 0 & \boldsymbol{B}_0 & \boldsymbol{B}_1 & \boldsymbol{B}_2 \end{bmatrix}. \tag{4.3}$$

We can now write the received N-vector signal $\boldsymbol{r}(k)$ in terms of the far-end N-vector signal $\boldsymbol{a}(k)$ as:

$$\boldsymbol{r}(k) = \boldsymbol{R}\_[\boldsymbol{H}\_[\boldsymbol{P}\_]_|]_|\boldsymbol{a}(k)_| \tag{4.4}$$

where the column-expanded input signal vector, $\boldsymbol{a}(k)_|$, is an $(n+m+l+1) \cdot N \times 1$ vector. Now let

$$\boldsymbol{V} = [\boldsymbol{H}\_[\boldsymbol{P}\_]_|]_| \tag{4.5}$$

which is an $N(l+1) \times N(m+n+l+1)$ convolution matrix representing the combined impulse response of the channel and transmitter filter. The input–output relation can now be written:

$$\boldsymbol{r}(k) = \boldsymbol{R}\_\boldsymbol{V}\boldsymbol{a}(k)_| + \boldsymbol{R}\_\boldsymbol{n}(k)_| \tag{4.6}$$

where $\boldsymbol{n}(k)_|$ is an additive noise signal at the output of the channel.

The receiver filter $\boldsymbol{R}$ is chosen to minimize the MSE:

$$E[\|\boldsymbol{r}(k) - \boldsymbol{a}(k-\phi)\|^2] \tag{4.7}$$

where ϕ represents the delay between the times when a symbol is transmitted and detected. Notice that $\boldsymbol{a}(k)$ is the N-dimensional data vector at time k, which is different

from the column-expanded vector $\boldsymbol{a}(k)_|$. Standard techniques show that $\boldsymbol{R}$ is determined by the familiar projection

$$\boldsymbol{R}\_ = \boldsymbol{E}_\phi^* \boldsymbol{V}^*(\boldsymbol{V}\Phi_a \boldsymbol{V}^* + \Phi_n)^{-1} \tag{4.8}$$

where $\Phi_a = E[\boldsymbol{a}(k)_|\boldsymbol{a}(k)_|^*]$, $\Phi_n = E[\boldsymbol{n}(k)_|\boldsymbol{n}(k)_|^*]$, and $\boldsymbol{E}_\phi = E[\boldsymbol{a}(k)_|\boldsymbol{a}^*(k - \phi)]$. If the data is uncorrelated with variance σ_d^2, then the $(m + n + l + 1) \cdot N \times N$ matrix $\boldsymbol{E}_\phi$ is given by:

$$\boldsymbol{E}_\phi = \sigma_d^2 \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \\ \cdots \\ \boldsymbol{I} \\ \cdots \\ \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix} \tag{4.9}$$

the $N \times N$ identity matrix being positioned after ϕ zero matrices, each $N \times N$. The integer ϕ can vary from 0 to $(m + n + l)$ and is selected to further minimize the MSE, which can be written as:

$$\text{MSE} = \text{trace}\ \{\phi_a - \boldsymbol{R}\_\boldsymbol{V}\boldsymbol{E}_\phi\} \tag{4.10}$$

where $\phi_a = E(\boldsymbol{a}(k)\,\boldsymbol{a}(k)^*]$.

Introducing near-end crosstalk adds a term to the expression for the received signal which depends on the near-end signal $\boldsymbol{b}(k)$:

$$\boldsymbol{r}(k) = \boldsymbol{R}\_\boldsymbol{V}\boldsymbol{a}(k)_| + \boldsymbol{R}\_\boldsymbol{G}\boldsymbol{b}(k)_| + \boldsymbol{R}\_\boldsymbol{n}(k)_| \tag{4.11}$$

where $\boldsymbol{G}$ is a convolution matrix incorporating the combined effect of the transmitter filter and crosstalk coupling impulse response, in analogy with $\boldsymbol{V}$. The same method then yields the MMSE solution:

$$\boldsymbol{R}\_ = \boldsymbol{E}_\phi^* \boldsymbol{V}^*(\boldsymbol{V}\Phi_a \boldsymbol{V}^* + \boldsymbol{G}\Phi_a \boldsymbol{G}^* + \Phi_n)^{-1} \tag{4.12}$$

with the same expression for minimum MSE given by (4.10).

We next consider the optimization of the transmitter filter given the receiver filter. The treatment is the same, except that we add a power constraint at the output of the transmitter filter. In analogy with $\boldsymbol{V}$, let the convolution matrix representing the combined impulse response of the channel and receiver filter be:

$$\boldsymbol{K} = [[\boldsymbol{R}_|]\_\boldsymbol{H}_|]\_ \tag{4.13}$$

and let the combined near-end crosstalk and receiver filter impulse response be denoted by the convolution matrix $\boldsymbol{S}$. Then the optimal transmitter filter is given by:

$$\boldsymbol{P}_| = (\boldsymbol{K}^*\boldsymbol{K} + \boldsymbol{S}^*\boldsymbol{S} + \mu\boldsymbol{I})^{-1}\boldsymbol{K}^*\boldsymbol{E}_{M-\phi} \tag{4.14}$$

where $M + 1$ is the number of data vectors in $\boldsymbol{a}(k)_|$, and μ is the Lagrange multiplier associated with the constraint on transmitter output power, and is determined by the condition:

$$\text{trace}\ \{\boldsymbol{P}_|\Phi_a \boldsymbol{P}_|^*\} = P_{\text{trans}}. \tag{4.15}$$

Note that the matrix $\boldsymbol{E}_{M-\phi}$ can be defined exactly as $\boldsymbol{E}_\phi$ where $\boldsymbol{a}(k)_|$ is replaced by the analogous column vector in which the data vectors are time reversed.

Finally, we mention that a decision feedback equalizer can be incorporated into the preceding optimizations using the methods in [13].

B. An Algorithm for the FIR Case

The problem of obtaining the simultaneously optimal pre- and post-filters in the FIR case in closed form is open. The following iterative method was found effective in practice: First optimize the receiver filter for fixed transmitter filter, then optimize the transmitter filter for that receiver filter, and repeat until convergence is obtained.

We observed that this iterative process can converge to different locally optimal solutions, depending on the choice of starting point, and in Section V we describe numerical results for some typical examples. We found in our simulations that the best results were obtained with the initial condition $\boldsymbol{P}_k = \boldsymbol{I}$ for $k = 0$, and $\boldsymbol{0}$ otherwise.

V. Numerical Results

A simplified model for two coupled twisted-pair wires is now used to compute optimal transmitter and receiver filters, and the associated MSE. We assume the model in Fig. 1(a), and that the receiver samples the channel output once per baud. The channel can therefore be replaced by an equivalent discrete-time channel, where the transmitter and receiver filters are discrete-time filters. Crosstalk in this case is equivalent to an additive WSS noise with spectrum that is proportional to the transmitted spectrum. The transmitter and receiver filters are assumed to be FIR, and our purpose is to study the MMSE as a function of sampling phase and how the matrix taps are allocated between the receiver and transmitter. This is in contrast to the numerical results presented in Section II-D, which illustrate the performance advantage obtained by the MMSE receiver, as compared with the type of receiver assumed here.

The channel model used here is essentially taken from [4]. In particular, the self-impulse response on each twisted pair was computed from a transmission line model for 12 kft of 24 gauge wire and 9 kft of 26 gauge wire. Each impulse response was sampled at 16 times the baud rate, which was chosen to be 800 kHz, and was truncated to 160 samples (10 symbol intervals), so that the number of matrix taps needed in the transmitter and receiver to substantially reduce the MSE would not be prohibitively large. It is assumed that $\sigma_n^2 = 0$, and that echo is perfectly canceled by echo cancelers, so that the diagonal elements of $\hat{\boldsymbol{G}}$ are zero.

Near-end crosstalk was generated by modifying the impulse response obtained from the transfer function [1]:

$$\hat{g}_{ik}(f) = j2\pi f \int_0^l e^{-2\Gamma(f)x} c_{ik}(x)\, dx \qquad (5.1)$$

where $c_{ik}(x)$ is the coupling capacitance between pairs i and k as a function of distance along the cable, l is the cable length, and $\Gamma(f) = \sqrt{(R + j2\pi fL)(G + j2\pi fC)}$ is the propagation constant where R, C, G, and L are the wire resistance, capacitance, conductance, and inductance of the cable per unit length, respectively. Here we assume for simplicity that $c_{ik}(x)$ is a constant independent of x, and that $c_{ik} = c_{ki}$, so that the matrix $\hat{\boldsymbol{G}}$ is symmetric. The near-end impulse response used to generate the numerical results was obtained by smoothing the impulse response obtained from this model. The smoothing was added to reduce the sensitivity of the MSE to the phase θ introduced in (2.1b). Far-end crosstalk was generated from the transfer function:

$$\hat{h}_{ik}(f) = j2\pi f e^{-2\Gamma(f)l} \int_0^l c_{ik}(x)\, dx. \qquad (5.2)$$

The transmitted signal is assumed to be of the form

$$\boldsymbol{a}(t) = \sum_k \boldsymbol{a}_k p(t - kT) \qquad (5.3)$$

where the $\boldsymbol{a}_k$'s are uncorrelated and have component ± 1, and $p(t)$ is a rectangular pulse shape of width T and amplitude one. The convolution of $p(t)$ with the channel impulse responses described in the preceding paragraph are shown in Figs. 5 and 6. The channel output is synchronously sampled at the baud rate, $1/T = 800$ kHz, so that the equivalent discrete-time model in Fig. 1(a) has as inputs the binary-valued vector input sequences $\{\boldsymbol{a}_k\}$ and $\{\boldsymbol{b}_k\}$. Since the channel (matrix) impulse response was sampled at 16 times the baud rate, this gives 16^2 possible combinations of sampling phases from which to choose. The particular combination that minimizes the MSE was obtained by exhaustive search.

The algorithm described in Section IV-B was used to compute simultaneously optimal transmitter and receiver filters. Fig. 7 shows plots of MSE versus iteration for cases in which the total number of (matrix) taps allocated between the transmitter and receiver is 15. In particular, curves are shown for one, two, five, and ten taps in the transmitter. The sampling phases in each case were chosen to minimize MSE after convergence. In each case, the MSE decreases rapidly during the first 200 iterations, but takes more than 1000 iterations to approach its asymptotic value.

Fig. 8 shows the improvement in MSE obtained by adding taps at the transmitter for a fixed-length receiver filter. Also shown is MSE versus the number of transmitter taps when the *total* number of transmitter and receiver taps is fixed. The results in Fig. 8 assume that the timing phase on each channel is chosen to minimize the MSE. For the channel model considered, the MSE is an insensitive function of how the taps are allocated between the transmitter and receiver. It is suspected that the variation shown (approximately 2 dB) is due to the quantization of sampling phase. In particular, the same results were gen-

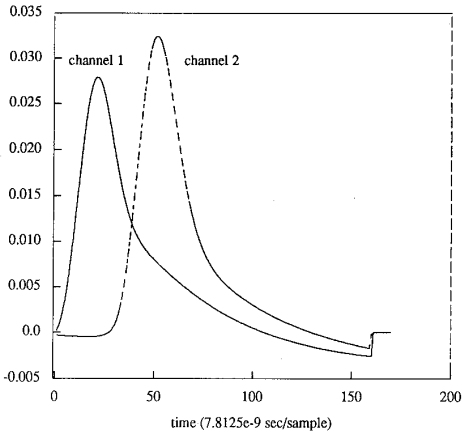


Fig. 5. Self-impulse responses of SISO channels.

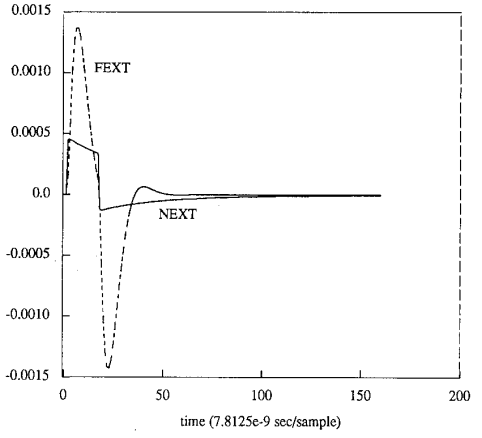


Fig. 6. NEXT and FEXT impulse responses.

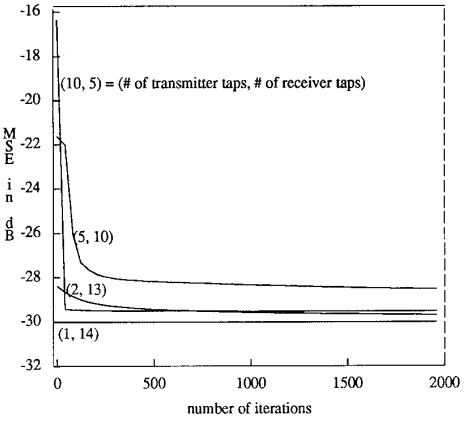


Fig. 7. MSE versus the number of iterations for different combinations of transmitter and receiver taps.

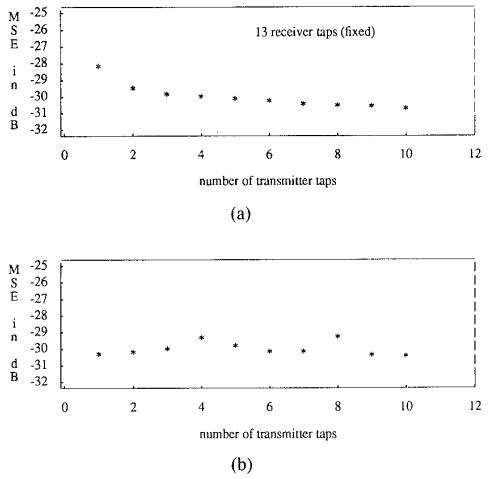


Fig. 8. MSE versus the number of transmitter taps assuming: (a) the number of receiver taps is held constant at 13; and (b) the total number of transmitter and receiver taps is 14.

erated assuming that the sampling phase is quantized to 8 samples per baud, and the variation in MSE was found to be much greater. The reason for this insensitivity is that as the transmitted power increases, so does near-end crosstalk so that, in the absence of noise, the MSE is not a decreasing function of transmitted power. For the case considered, where near-end crosstalk is the dominant impairment, the transmitted power constraint (4.15) therefore becomes extraneous. Consequently, the MSE becomes almost independent of where filters are added relative to the channel.

The dependence of MSE on timing phase is illustrated in Fig. 9, which shows a two-dimensional perspective plot of MSE as a function of timing phase on both channels for the case where the receiver filter has 14 taps and the transmitter filter has 1 tap. The variation between the minimum and maximum values of MSE in this figure is approximately 22 dB. For the case considered, the MSE was found to be a very sensitive function of timing phase. However, as in the SISO case, as taps are added to the receiver filter, the dependence of the MSE on sampling phase diminishes. In principle, the phase θ in (2.1b), which is the difference between the times when near-end symbols are transmitted and far-end symbols are detected, can also be adapted to minimize MSE. However, because of the practical difficulties involved in such an optimization, this was not investigated.

VI. Conclusions

The MMSE linear equalizer in the presence of synchronous crosstalk has been completely specified in terms of the near- and far-end channel transfer functions. In practice, the tapped delay lines associated with the MMSE equalizer can be approximated by FIR filters, and the taps can be adapted to minimize the MSE between transmitted symbols and the filter output. In order to effectively suppress N different near-end crosstalk signals, the bandwidth available must be $N + 1$ times the Nyquist frequency. If the MMSE equalizer is implemented as a fractionally spaced equalizer, it follows that the tap spacing should be no greater than $T/(N + 1)$ to suppress all crosstalk signals.

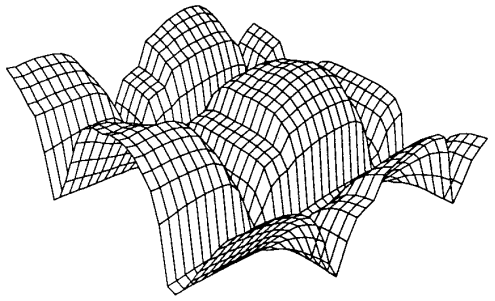


Fig. 9. Two-dimensional perspective plot of MSE versus timing phases on each channel.

Optimization of the transmitter filter in the presence of crosstalk, assuming either channel model in Fig. 1, appears to be difficult since the effect of aliasing must be taken into account. However, if the channel is bandlimited to the Nyquist interval, a closed-form solution exists for the MMSE transmitter filter in particular cases, since crosstalk at the sampling intervals can be treated as an additive WSS noise.

Our numerical results in the preceding section demonstrate that when near-end crosstalk is the dominant channel impairment, adding taps to an MIMO transmitter filter selected to minimize MSE has the same effect as adding taps to a MIMO linear equalizer. From a practical point of view, it is therefore better to equalize the channel at the receiver since, in the case of an unknown or time-varying channel, the equalizer taps can be adapted to minimize the MSE. In contrast, an adaptive algorithm for the transmitter taps requires an estimate of the channel and crosstalk impulse responses, which can be provided only by feedback channels. For the channel model considered, the MSE was found to be a sensitive function of timing phase, which was selected independently on both SISO channels.

A natural extension of this work is to consider more sophisticated detection algorithms in the presence of crosstalk, such as decision feedback equalization, and maximum-likelihood estimation. Decision feedback equalization in the presence of crosstalk is considered in [10], [13], [17], and [18]. For a 2×2 MIMO channel, it has been demonstrated in [13] that sensitivity to timing phase can be reduced by using an MIMO decision feedback structure, which has the approximate effect of shortening the channel impulse response. For the type of crosstalk channels considered here, simulation results in [18] show that timing sensitivity can also be significantly reduced by using a fractionally spaced equalizer. Maximum-likelihood and related estimation techniques are considered in [33]–[35] for the multiuser channel. The

performance of these techniques for the types of channels considered here has not yet been studied.

Appendix A
Derivation of the MMSE Linear Equalizer

Consider the channel model in Fig. 2 with $M \times N$ transfer function $\hat{\boldsymbol{J}}$ and input $\boldsymbol{u}(t) = \Sigma_k \boldsymbol{a}_k p(t - kT)$, where $\{\boldsymbol{a}_k\}$ is the input N-vector data sequence. Suppose that this channel is followed by an $N \times M$ receiver filter $\hat{\boldsymbol{R}}$, and let $\boldsymbol{r}(0)$ denote the output of $\hat{\boldsymbol{R}}$ at time $t = 0$. Then the MSE can be written as:

$$\begin{aligned}\text{MSE} &= E\{\|\boldsymbol{r}(0) - \boldsymbol{a}_0\|^2\} \\ &= \text{trace}\left\{\boldsymbol{\phi}_0^d - 2\,\text{Re}\sum_k \boldsymbol{K}_k \boldsymbol{\phi}_k^d + \sum_k \sum_m \boldsymbol{K}_k^* \boldsymbol{K}_m \boldsymbol{\phi}_{k-m}^d \right. \\ &\quad \left. + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \boldsymbol{R}(s)\boldsymbol{\phi}_n(s-u)\boldsymbol{R}^*(u)\,ds\,du\right\} \quad \text{(A.1)}\end{aligned}$$

where $\boldsymbol{R}$ is the impulse response of $\hat{\boldsymbol{R}}$, $\boldsymbol{K}_k = \boldsymbol{R} * \tilde{\boldsymbol{J}}(kT)$, $\boldsymbol{\phi}_k^d = E(\boldsymbol{a}_m \boldsymbol{a}_{m+k}^*)$, $\boldsymbol{\phi}_n(s) = E[\boldsymbol{n}(t)\boldsymbol{n}^*(t-s)]$, and $\tilde{\boldsymbol{J}}(t) = \boldsymbol{J} * \boldsymbol{P} * p(t)$. In general $\tilde{\boldsymbol{J}}(t)$ is an $M \times N$ matrix, $\boldsymbol{R}$ is an $N \times M$ matrix, $\boldsymbol{\phi}_n$ is an $M \times M$ matrix, and $\boldsymbol{\phi}_k^d$ and $\boldsymbol{K}_k$ are $N \times N$ matrices. A standard variational argument can be used to derive the following optimality condition for the receiver filter

$$\int_{-\infty}^{\infty} \boldsymbol{R}(s)\boldsymbol{\phi}_n(t-s)\,ds = \sum_k (\boldsymbol{\phi}_k^d - \tilde{\boldsymbol{K}}_k)\tilde{\boldsymbol{J}}^*(kT - t) \quad \text{(A.2)}$$

where

$$\tilde{\boldsymbol{K}}_k = \sum_m \boldsymbol{K}_m \boldsymbol{\phi}_{k-m}^d. \quad \text{(A.3)}$$

Taking the Fourier transform of both sides of (A.2) gives:

$$\hat{\boldsymbol{R}}\boldsymbol{S}_n = \left[\sum_k (\boldsymbol{\phi}_k^d - \tilde{\boldsymbol{K}}_k)e^{-j2\pi fkT}\right]\hat{\tilde{\boldsymbol{J}}}^* \quad \text{(A.4)}$$

where $\hat{\tilde{\boldsymbol{J}}}$ is the Fourier transform of $\tilde{\boldsymbol{J}}$.

Defining the Fourier series associated with the sequence $\{\boldsymbol{K}_k\}$ as:

$$\hat{\boldsymbol{K}}(e^{j2\pi f}) = \sum_k \boldsymbol{K}_k e^{-j2\pi fkT} \quad \text{(A.5)}$$

and the Fourier series $\hat{\tilde{\boldsymbol{K}}}(e^{j2\pi f})$ in an analogous manner, it then follows from (A.3) that:

$$\hat{\tilde{\boldsymbol{K}}}(e^{j2\pi f}) = \hat{\boldsymbol{K}}(e^{j2\pi f})\boldsymbol{S}^d(e^{j2\pi f}) \quad \text{(A.6)}$$

where $\boldsymbol{S}^d(e^{j2\pi f})$ is the data spectrum defined by (2.6). Consequently, (A.4) can be written as:

$$\begin{aligned}\hat{\boldsymbol{R}}(f)\boldsymbol{S}_n(f) &= (\boldsymbol{I} - \hat{\boldsymbol{K}}(e^{j2\pi f}))\boldsymbol{S}^d(e^{j2\pi f})\hat{\tilde{\boldsymbol{J}}}^*(f) \\ &= \hat{\boldsymbol{Q}}(e^{j2\pi f})\hat{\tilde{\boldsymbol{J}}}^*(f) \quad \text{(A.7)}\end{aligned}$$

which corresponds to the block diagram in Fig. 2, where $\hat{\boldsymbol{Q}}(e^{j2\pi f}) = [\boldsymbol{I} - \hat{\boldsymbol{K}}(e^{j2\pi f})]\boldsymbol{S}^d(e^{j2\pi f})$.

We now wish to find an expression for the transfer function of the tapped delay line $\hat{\boldsymbol{Q}}(e^{j2\pi f})$ in terms of the channel transfer function $\hat{\boldsymbol{J}}$, and the data and noise spectra. Let $\{\boldsymbol{x}_k\}$ denote the input sequence to the tapped delay line at the sampling times kT, $k = 0, 1, 2, \cdots$, and let $\boldsymbol{X}(z)$ denote the associated z-transform. If $\boldsymbol{A}(z)$ denotes the z-transform of the data sequence and $\boldsymbol{N}(z)$ denotes the z-transform of the sampled output of the matched filter with only the noise $\boldsymbol{n}(t)$ as the input, then:

$$\boldsymbol{X}(z) = \boldsymbol{S}^J(z)\boldsymbol{A}(z) + \boldsymbol{N}(z) \quad \text{(A.8)}$$

where $\boldsymbol{S}^J(z)$ is the z-transform of the equivalent discrete-time channel with inputs $\{\boldsymbol{a}_k\}$ and outputs $\{\boldsymbol{x}_k\}$. It follows that $\boldsymbol{S}^J(z)$ for $z = e^{j2\pi f}$ is given by (2.5), and the spectrum of the noise sequence at the output of the matched filter is $\boldsymbol{S}^N(e^{j2\pi f}) = \boldsymbol{S}^J(e^{j2\pi f})$. (See [30, Ch. 3]. The generalization to MIMO systems is straightforward.) Let $\{\boldsymbol{y}_k\}$ be the sequence of outputs of $\hat{\boldsymbol{Q}}(z)$, that is, $\boldsymbol{y}_k = \Sigma_i \boldsymbol{Q}_{k-i}\boldsymbol{x}_i$, where $\{\boldsymbol{Q}_i\}$ is the impulse response of the tapped delay line in Fig. 2. $\boldsymbol{y}_k$ is, therefore, the MMSE estimate of $\boldsymbol{a}_k$. The z-transform of the error sequence $\{\boldsymbol{e}_k = \boldsymbol{y}_k - \boldsymbol{a}_k\}$ is:

$$\boldsymbol{E}(z) = [\hat{\boldsymbol{Q}}(z)\boldsymbol{S}^J(z) - \boldsymbol{I}]\boldsymbol{A}(z) + \boldsymbol{N}(z) \quad \text{(A.9)}$$

so that the MSE can be written as:

$$\begin{aligned}\text{MSE} = \text{trace}\left\{\frac{1}{2\pi j}\oint \left([\hat{\boldsymbol{Q}}(z)\boldsymbol{S}^J(z) - \boldsymbol{I}]\right.\right. \\ \cdot \boldsymbol{S}^d(z)[\boldsymbol{S}^J(z)\hat{\boldsymbol{Q}}'(z^{-1}) - \boldsymbol{I}] \\ \left.\left. + \hat{\boldsymbol{Q}}(z)\boldsymbol{S}^N(z)\hat{\boldsymbol{Q}}'(z^{-1})\right)\frac{dz}{z}\right\} \quad \text{(A.10)}\end{aligned}$$

where prime denotes transpose. (Note that $\boldsymbol{S}^J(z) = [\boldsymbol{S}^J(z^{-1})]'$.) Selecting $\hat{\boldsymbol{Q}}(z)$ to minimize this expression gives (2.4). Combining (A.1), (A.2), and (A.6) gives the following expression for the MMSE:

$$\begin{aligned}\text{MMSE} &= \text{trace}\,[(\boldsymbol{\phi}_0^d - \tilde{\boldsymbol{K}}_0) \\ &= \text{trace}\left\{T\int_{-1/(2T)}^{1/(2T)} [\boldsymbol{I} - \hat{\boldsymbol{K}}(e^{j2\pi f})]\boldsymbol{S}_d(e^{j2\pi f})\,df\right\} \\ &= \text{trace}\left\{T\int_{-1/(2T)}^{1/(2T)} \hat{\boldsymbol{Q}}(e^{j2\pi f})\,df\right\}. \quad \text{(A.11)}\end{aligned}$$

Appendix B
Derivation of (3.11)

Substituting (3.5) into (3.10) gives:

$$\begin{aligned}&(\hat{\boldsymbol{H}}^*\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{H}}\hat{\boldsymbol{P}}\boldsymbol{S}_d^2\hat{\boldsymbol{P}}^*\hat{\boldsymbol{H}}^*\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{H}} \\ &\quad + \hat{\boldsymbol{G}}^*\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{H}}\hat{\boldsymbol{P}}\boldsymbol{S}_d^2\hat{\boldsymbol{P}}^*\hat{\boldsymbol{H}}^*\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{G}} + \mu\boldsymbol{I})\hat{\boldsymbol{P}} \\ &\quad = \hat{\boldsymbol{H}}^*\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{H}}\hat{\boldsymbol{P}}\boldsymbol{S}_d \quad \text{(B.1)}\end{aligned}$$

where $\hat{\boldsymbol{A}}$ is defined by (3.6), and μ is the Lagrange multiplier chosen to satisfy (3.9). Letting $\boldsymbol{S}_d = \sigma_d^2\boldsymbol{I}$, and post-multiplying by $\hat{\boldsymbol{P}}^*$ and then by $\hat{\boldsymbol{W}}^{-1}$ gives:

$$\begin{aligned}&\hat{\boldsymbol{H}}^*\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{H}} - \sigma_d^2\hat{\boldsymbol{H}}^*\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{H}}\hat{\boldsymbol{W}}\hat{\boldsymbol{H}}^*\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{H}} \\ &\quad - \sigma_d^2\hat{\boldsymbol{H}}^*\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{H}}\hat{\boldsymbol{W}}\hat{\boldsymbol{G}}^*\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{G}} = \frac{\mu}{\sigma_d^2}\boldsymbol{I}. \quad \text{(B.2)}\end{aligned}$$

Now

$$\begin{aligned}\hat{\boldsymbol{H}}^*\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{H}} - \sigma_d^2\hat{\boldsymbol{H}}^*\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{H}}\hat{\boldsymbol{W}}\hat{\boldsymbol{H}}^*\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{H}} \\ = \hat{\boldsymbol{H}}^*(\hat{\boldsymbol{A}}^{-1} - \sigma_d^2\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{H}}\hat{\boldsymbol{W}}\hat{\boldsymbol{H}}^*\hat{\boldsymbol{A}}^{-1})\hat{\boldsymbol{H}} \\ = \hat{\boldsymbol{H}}^*(\hat{\boldsymbol{A}}^{-1}(\hat{\boldsymbol{A}} - \sigma_d^2\hat{\boldsymbol{H}}\hat{\boldsymbol{W}}\hat{\boldsymbol{H}}^*)\hat{\boldsymbol{A}}^{-1})\hat{\boldsymbol{H}} \\ = \hat{\boldsymbol{H}}^*\hat{\boldsymbol{A}}^{-1}(\sigma_d^2\hat{\boldsymbol{G}}\hat{\boldsymbol{W}}\hat{\boldsymbol{G}}^* + \boldsymbol{S}_n)\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{H}} \qquad \text{(B.3)}\end{aligned}$$

and combining (B.3) with (B.2) gives (3.11).

To show that $\hat{\boldsymbol{W}}$ given by (3.15) satisfies (3.14), we first rewrite (3.14) as:

$$\frac{1}{\xi\mu}\hat{\boldsymbol{H}}\hat{\boldsymbol{H}}^* = \hat{\boldsymbol{H}}\hat{\boldsymbol{W}}\hat{\boldsymbol{H}}^*\hat{\boldsymbol{H}}\hat{\boldsymbol{W}}\hat{\boldsymbol{H}}^* + \frac{2}{\xi}\hat{\boldsymbol{H}}\hat{\boldsymbol{W}}\hat{\boldsymbol{H}}^* + \frac{1}{\xi^2}\boldsymbol{I}. \qquad \text{(B.4)}$$

Postmultiplying by $\hat{\boldsymbol{H}}$ and then $(\hat{\boldsymbol{H}}^*\hat{\boldsymbol{H}})^{-1}$ gives:

$$\frac{1}{\xi\mu}\hat{\boldsymbol{H}} = \hat{\boldsymbol{H}}\hat{\boldsymbol{W}}\hat{\boldsymbol{H}}^*\hat{\boldsymbol{H}}\hat{\boldsymbol{W}} + \frac{2}{\xi}\hat{\boldsymbol{H}}\hat{\boldsymbol{W}} + \frac{1}{\xi^2}\hat{\boldsymbol{H}}(\hat{\boldsymbol{H}}^*\hat{\boldsymbol{H}})^{-1} \qquad \text{(B.5)}$$

and premultiplying by $\hat{\boldsymbol{H}}^*$ and then $(\hat{\boldsymbol{H}}^*\hat{\boldsymbol{H}})^{-1}$ gives

$$\frac{1}{\xi\mu}\boldsymbol{I} = \left(\hat{\boldsymbol{W}}\hat{\boldsymbol{H}}^*\hat{\boldsymbol{H}} + \frac{2}{\xi}\boldsymbol{I}\right)\hat{\boldsymbol{W}} + \frac{1}{\xi^2}(\hat{\boldsymbol{H}}^*\hat{\boldsymbol{H}})^{-1}. \qquad \text{(B.6)}$$

It is now easily verified that the expression for $\hat{\boldsymbol{W}}$ given by (3.15) satisfies (B.6).

We now show that (3.15) is the same as the MMSE transmitter transfer function obtained in [12] and [22]. Since $\hat{\boldsymbol{H}}^*\hat{\boldsymbol{H}}$ is real and symmetric, it can be rewritten as $\Psi(f)\Lambda(f)\Psi^*(f)$, where Ψ is orthonormal for each f and Λ is diagonal. Let $\tilde{\boldsymbol{P}} = \Psi\hat{\boldsymbol{P}}$ and $\hat{\boldsymbol{T}} = \tilde{\boldsymbol{P}}\tilde{\boldsymbol{P}}^*$. It is shown in [12] and [22] that the value of $\hat{\boldsymbol{T}}$ which minimizes the MSE is:

$$\hat{\boldsymbol{T}} = \frac{1}{\sqrt{\xi\mu}}\Lambda^{-1/2} - \frac{1}{\xi}\Lambda^{-1} \qquad \text{(B.7)}$$

so that

$$\begin{aligned}\hat{\boldsymbol{W}} = \Psi^*\hat{\boldsymbol{T}}\Psi = \frac{1}{\sqrt{\mu\xi}}\Psi^*\Lambda^{-1/2}\Psi - \frac{1}{\xi}\Psi^*\Lambda^{-1}\Psi \\ = \frac{1}{\sqrt{\mu\xi}}(\hat{\boldsymbol{H}}^*\hat{\boldsymbol{H}})^{-1/2} - \frac{1}{\xi}(\hat{\boldsymbol{H}}^*\hat{\boldsymbol{H}})^{-1}. \qquad \text{(B.8)}\end{aligned}$$

ACKNOWLEDGMENT

The authors thank D. D. Falconer and B. Petersen for discussion and comments, which helped improve both the original and revised manuscripts, and for bringing many of the cited references on MMSE equalization to their attention. The authors also thank G. Golden for helpful comments and discussions concerning this work.

REFERENCES

[1] H. Cravis and T. V. Crater, "Engineering of T1 carrier system repeatered lines," *Bell Syst. Tech. J.*, vol. 42, no. 2, pp. 431–486, Mar. 1963.

[2] A. R. Kaye and D. A. George, "Transmission of multiplexed PAM signals over multiple channel and diversity systems," *IEEE Trans. Commun. Techn.*, vol. COM-18, no. 5, pp. 520–526, Oct. 1970.

[3] M. L. Honig, K. Steiglitz, and B. Gopinath, "Multi-channel signal processing fror data communications in the presence of crosstalk," *IEEE Trans. Commun.*, vol. 38, no. 4, pp. 551–558, Apr. 1990.

[4] P. Crespo and M. L. Honig, "A simulation study of near- and far-end crosstalk cancellation for multi-channel data transmission," in *Advances in Communications and Signal Processing*, W. A. Porter and S. C. Kak, Eds. New York: Springer-Verlag, 1989, pp. 219–230.

[5] S. Verdú, "Recent progress in multi-user detection," in *Advances in Communications and Signal Processing*, W. A. Porter and S. C. Kak, Eds. New York: Springer-Verlag, 1989, pp. 27–38.

[6] K. S. Schneider, "Crosstalk resistant receiver for *M*-ary multiplexed communications," *IEEE Trans. Aerospace and Electron. Syst.*, vol. AES-16, no. 4, pp. 426–433, July 1980.

[7] W. van Etten, "An optimum linear receiver for multiple channel digital transmission systems," *IEEE Trans. Commun.*, pp. 828–834, Aug. 1975.

[8] D. Shnidman, "A generalized Nyquist criterion and an optimum linear receiver for a pulse modulation system," *Bell Syst. Tech. J.*, pp. 2163–2177, Nov. 1967.

[9] P. Monsen, "MMSE equalization of interference on fading diversity channels," *IEEE Trans. Commun.*, vol. COM-32, no. 1, pp. 5–12, Jan. 1984.

[10] D. G. Messerschmitt and L. C. Barbosa, "A study of sampling detectors for magnetic recording," IBM Res. Rep. RJ 4081, July 1984.

[11] J. Cioffi *et al.*, "Adaptive equalization in magnetic-disk storage channels," *IEEE Commun. Mag.*, vol. 28, no. 2, pp. 14–29, Feb. 1990.

[12] N. Amitay and J. Salz, "Linear equalization theory in digital data transmission over dually polarized fading radio channels," *Bell Syst. Tech. J.*, vol. 63, no. 10, pp. 2215–59, Dec. 1984.

[13] M. Kavehrad and J. Salz, "Cross-polarization cancellation and equalization in digital transmission over dually polarized multipath fading channel," *Bell Syst. Tech. J.*, vol. 64, no. 10, pp. 2211–45, Dec. 1985.

[14] D. C. Cox, "Universal digital portable radio communications," *Proc. IEEE*, vol. 75, no. 4, pp. 436–477, Apr. 1987.

[15] J. C. Campbell, A. J. Gibbs, and B. M. Smith, "The cyclostationary nature of crosstalk interference from digital signals in multipair cable—Part I: Fundamentals, and Part II: Applications and further results," *IEEE Trans. Commun.*, vol. COM-31, no. 5, pp. 629–649, May 1983.

[16] W. A. Gardner and L. E. Franks, "Characterization of cyclostationary random signal processes," *IEEE Trans. Inform. Theory*, vol. IT-21, no. 1, pp. 4–14, Jan. 1975.

[17] B. R. Petersen and D. D. Falconer, "Equalization in cyclostationary interference," Rep. SCE-90-01, Dept. Syst. Comput. Eng., Carleton Univ., Ottawa, Ont., Canada, Jan. 1990.

[18] M. Abdulrahman and D. D. Falconer, "Cyclostationary crosstalk suppression by decision feedback equalization on digital subscriber loops," *IEEE J. Select Areas Commun.*, this issue, pp. 640–649.

[19] D. J. Harrison, "Adaptive equalization for channels with crosstalk," M. Eng. Thesis, Carleton Univ., Ottawa, Ont., Canada, Dec. 1969.

[20] E. Biglieri, M. Elia, and L. Lopresti, "The optimal linear receiving filter for digital transmission over nonlinear channels," *IEEE Trans. Inform. Theory*, vol. 35, no. 3, pp. 620–625, May 1989.

[21] P. Crespo, M. L. Honig, and K. Steiglitz, "Optimization of pre- and post filters in the presence of near- and far-end crosstalk," presented at 1989 IEEE Int. Conf. Commun., Boston, MA.

[22] J. Salz, "Digital transmission over cross-coupled linear channels," *Bell Syst. Tech. J.*, vol. 64, no. 6, pp. 1147–1159, July/Aug. 1985.

[23] A. Duel-Hallen, "Equalizers for multiple input/multiple output channels and PAM systems with cyclostationary input sequences," *IEEE J. Select. Areas Commun.*, this issue, pp. 630–639.

[24] K.-H. Lee and D. P. Petersen, "Optimal linear coding for vector channels," *IEEE Trans. Commun.*, vol. COM-24, no. 12, 1288–1290, Dec. 1976.

[25] H. S. Malvar and D. H. Staelin, "Optimal pre and post-filters for multichannel signal processing," *IEEE Trans. Sig. Proces.*, vol. 36, no. 2, pp. 287–289, Feb. 1988.

[26] H. P. Kramer and M. V. Mathews, "A linear coding for transmitting a set of correlated signals," *IRE Trans. Inform. Theory*, vol. IT-2, pp. 41–46, Sept. 1956.

[27] J. W. Lechleider and K. Sistanizadeh, "Two design techniques for transceivers for digital subscriber lines," in *Proc. 1988 GLOBECOM Conf.*, Dec. 1988, paper 35.6.1, pp. 1150–1154.

[28] D. D. Falconer and B. R. Petersen, private correspondence.

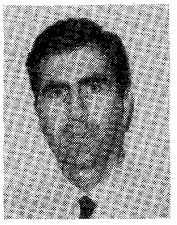
[29] G. Golden, private correspondence.

[30] E. A. Lee and D. G. Messerschmitt, *Digital Communications*. Boston, MA: Kluwer, 1988.
[31] S. Qureshi, "Adaptive equalization," *Proc. IEEE*, vol. 73, no. 9, pp. 1349–1387, Sept. 1985.
[32] S. Benedetto, E. Biglieri, and V. Castellani, *Digital Transmission Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1987.
[33] W. van Etten, "Maximum likelihood receiver for multiple channel transmission systems," *IEEE Trans. Commun.*, pp. 276–283, Feb. 1976.
[34] H. V. Poor and S. Verdú, "Single-user detectors for multiuser channels," *IEEE Trans. Commun.*, vol. 36, no. 1, pp. 50–60, Jan. 1988.
[35] S. Verdú, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol. IT-32, no. 1, pp. 85–96, Jan. 1986.
[36] W. A. Gardner, "On the optimal linear receiving filter for digital transmission over nonlinear channels," *IEEE Trans. Inform. Theory*, vol. 37, no. 1, p. 219, Jan. 1991.

Michael L. Honig was born in Phoenix, AZ, in 1955. He received the B.S. degree in electrical engineering from Stanford, University, Stanford, CA, in 1977, and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1978 and 1981, respectively.

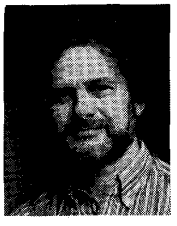
From July 1981 to October 1983, he was a Member of the Technical Staff at AT&T Information Systems, formerly part of Bell Laboratories, Holmdel, NJ, where he worked on the design and performance analysis of local area networks and on voiceband data transmission. He subsequently transferred to the Systems Principles Research Division at Bell Communications Research, where he is currently working in the areas of data communications and signal processing.

Dr. Honig is a freelance trombonist, a member of Tau Beta Pi and Phi Beta Kappa, and is currently an Associate Editor for the *IEEE Transactions on Communications*.



Pedro Crespo was born in Barcelona, Spain, in 1955. He received the engineering degree from E.T. Superior de Ingenieros de Telecommunicacion, Barcelona, Spain, in 1978 and the M.S. and Ph.D. degrees in electrical engineering from the University of Southern California, Los Angeles, in 1981 and 1984, respectively.

From September 1984 to April 1990, he was a Member of the Technical Staff at Bell Communications Research, Morristown, NJ, where he investigated problems in data communications and signal processing. At present, he is a District Manager at Telefonica I+D, Madrid, Spain.



Kenneth Steiglitz (S'57–M'64–SM'79–F'81) was born in Weehawken, NJ, in January 1939. He received the B.E.E. (magna cum laude), M.E.E., and Eng.Sc.D. degrees from New York University, New York, NY, in 1959, 1960, and 1963, respectively.

Since September 1963, he has been at Princeton University, Princeton, NJ, where he is now Professor of Computer Science, teaching and conducting research on VLSI design and implementation of signal processing, optimization algorithms, and the foundations of computing. He is the author of *Introduction to Discrete Systems* (New York: Wiley, 1974) and coauthor, with C. H. Papadimitriou, of *Combinatorial Optimization: Algorithms and Complexity* (Englewood Cliffs, NJ: Prentice-Hall, 1982).

Dr. Steiglitz is a member of the VLSI Committee of the IEEE ASSP Society, is serving his second term as member of the Administrative Committee, and has also served on the Digital Signal Processing Committee, as Awards Chairman of that Society, and as chairman of their Technical Direction Committee. He is an Associate Editor of the journal *Networks*, and is a former Associate Editor of the *Journal of the Association for Computing Machinery*. A member of Eta Kappa Nu, Tau Beta Pi, and Sigma Xi, he received the Technical Achievement Award of the ASSP Society in 1981, their Society Award in 1986, and the IEEE Centennial Medal in 1984.