Minimum Mean-Squared Error Echo Cancellation and Equalization for Digital Subscriber Line Transmission: Part I—Theory and Computation

DAVID W. LIN, SENIOR MEMBER, IEEE

Abstract—**An integral part of ISDN (integrated services digital network) is the provision of full-duplex digital transmission capability over voice-grade metallic subscriber lines with all the associated disturbances from echo, intersymbol interference, and crosstalk, among other things. We here present a theory for analyzing the optimal performance, in MMSE (minimum mean-squared error) sense, of full-duplex transceiver structures incorporating echo cancellers and decision-feedback equalizers. This theory augments previously published results by allowing a colored input signal and a fractionally spaced multitap MMSE forward filter in the decision-feedback equalizer. Computational considerations of the theoretical results are also addressed, where we investigate the properties of and efficient ways of calculating the MMSE solutions for various types of line codes, including the precoded partial-response codes and the block codes. We also discuss on how the MMSE is related to the ubiquitous SNR (signal-to-noise ratio) measure and the concerns associated with using it to gauge the transmission performance.**

I. Introduction

AN INTEGRAL part of ISDN (integrated services digital network) is the provision of full-duplex digital transmission capability over the pair of metal wires that connects a subscriber to a telephone company's terminal equipment where these metallic subscriber lines were originally designed for voice communications. Subscriber lines with digital transmission capability have been termed "digital subscriber lines," or DSL in short. Full-duplex DSL transceivers have to deal with echo, intersymbol interference (ISI), and crosstalk from nearby transmission paths, among other things. Fig. 1 shows a skeleton DSL transceiver structure along with the environment it operates in. Several key components in the transceiver are the line coder, the echo canceler, and the equalizer. In our study, we assume that the scrambler generates an i.i.d. (independent and identically distributed) binary sequence. We are interested in analyzing the optimal performance, in the MMSE (minimum mean-squared error) sense, of such a transceiver. The MMSE criterion has been widely used in practice because of its simplicity and the existence of simple adaptive filtering algorithms which converge to the MMSE solution. Further, the MMSE is often related to the ubiquitous SNR (signal-to-noise ratio) performance measure in a very simple way.

Despite the amount of literature on echo cancellation and equalization [1], [2], we found that the existing theory needed a fair amount of augmentation in two aspects to study the problem at hand. First, the existing theory largely handles i.i.d. signals only, while there has been a long standing interest in modulation and line coding schemes which yield non-i.i.d. and even correlated symbol sequences. A most recent example of such interest is seen in the standardization effort toward a DSL transmission scheme for ISDN Basic Access in the United States—there a number of line codes which generate non-i.i.d. symbol sequences were proposed. Secondly, for the equalizer alone, a much-studied structure is the decision-feedback with a multitap forward filter. Nevertheless, previous results on combined echo cancelation and equalization, represented by those of Mueller [3] and Falconer [4], only considered single-tap forward filters. However, Mueller [3] and Falconer [4] also investigated the convergence behaviors of the jointly adaptive echo canceller and equalizer; while we shall only be concerned with the optimal performance in the MMSE sense.

Fig. 2 shows the internals of the (decision-feedback) equalizer and its relative position with the echo canceller. Jointly, the EC (echo canceller) and the DFE (decision-feedback equalizer) seek to minimize $E\{\epsilon^2\}$. A different arrangement, alluded to in [5] as a possibility, is shown in Fig. 3. Here the EC and the DFE are independently adapted to minimize $E\{\delta^2\}$ and $E\{\epsilon^2\}$, respectively. The problem with this latter arrangement is that, for a given EC length, the optimal EC (in the sense of minimizing $E\{\delta^2\}$) need not be one which minimizes $E\{\epsilon^2\}$. As a result, lengthening the MMSE EC (which reduces the minimum mean-squared δ) does not necessarily reduce the minimum mean-squared decision-point error ϵ. In some cases, the contrary is true. This undesirable phenomenon has been observed in our simulation study. We will give a pedagogical example to illustrate the point.

Note that the above does not establish any superiority of the jointly adapted structure in terms of yielding a lower minimum mean-squared ϵ. Whether this will be the case depends on channel and echo path characteristics. The above only points out a potential unhappy surprise that may lie in the way of the separately adapted structure. More will be said about this when we come to the pedagogical example.

This problem, of course, does not exist if the echo is fully cancelled.

In Section II, we derive general equations to calculate the MMSE results for both the structures of Figs. 2 and 3. We also assume that the only significant sources of noise are echo, ISI, and near-end crosstalk (NEXT) from identical DSL transmission systems (self-NEXT, in short). In principle, it is easy to consider other noise sources as well. But they are assumed negligible for this study.

For some line codes, a direct application of the MMSE equations derived in Section II can be computationally very intensive. We thus devote a major part of Section III to address this issue. Several means of speeding up the computation in various cases are outlined, as they are not completely obvious. Some results are interesting by themselves, apart from contributing to computational saving. In Section IV, we discuss on how the MMSE is related to the SNR measure and the concerns associated with using it to gauge the DSL performance.

A companion paper [22] reports and discusses some simulation results. Space constraint prevents us from considering here the important limiting case of an infinitely long DFE, which poses interest-

Paper approved by the Editor for Channel Equalization of the IEEE Communications Society. Manuscript received November 24, 1987; revised July 4, 1988.

The author is with Bell Communications Research, Red Bank, NJ 07701.

IEEE Log Number 8932071.

0090-6778/90/0100-0031$01.00

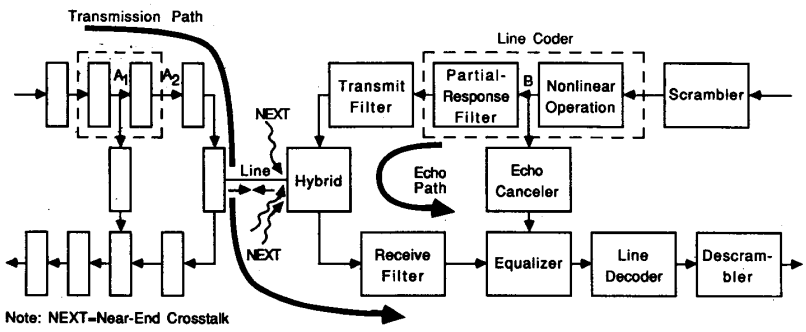

Fig. 1. A skeleton DSL transmission system.

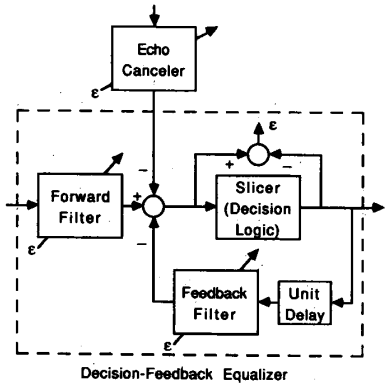

Fig. 2. Jointly adapted echo canceller and decision-feedback equalizer.

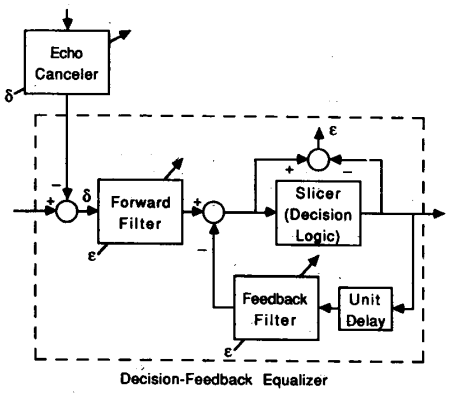

Fig. 3. Separately adapted echo canceller and decision-feedback equalizer.

ing theoretical and computational problems and warrants a separate treatise [23].

II. MMSE Echo Cancellation and Equalization

A. The Jointly Adapted Structure

Fig. 4 shows a conceptualized DSL transmission path for the environment of Fig. 1 and the transceiver structure of Fig. 2. In the figure, the "channel" includes all analog transmission filtering. For generality, the DFE forward filter is assumed to be (possibly) fractionally spaced with a temporal tap spacing kT/m where T is the symbol period and k and m are coprime integers with $k < m$, except in the case of a synchronous equalizer where $k = m = 1$. The EC and the DFE feedback filter are assumed to have a tap spacing T. $\boldsymbol{h}$ denotes the vector of channel impulse response sampled at rate m/T, and Θ the corresponding *folded* frequency response. The NEXT and the echo are also assumed to be sampled at rate T/m. $\boldsymbol{h}_e$, $\boldsymbol{h}_g$, $\boldsymbol{h}_f$, and $\boldsymbol{h}_b$ denote, respectively, the impulse response vectors of the echo path, the echo canceler, the DFE forward filter, and the DFE feedback filter; while E, G, F, and B their frequency responses. Note that the tap spacing is T/m in $\boldsymbol{h}_e$, k/mT in $\boldsymbol{h}_f$, and T in $\boldsymbol{h}_g$ and $\boldsymbol{h}_b$.

Assume that the slicer in the DFE makes no decision error. Assume also that the cursor delay is c symbol periods. Then the transmission path is mathematically equivalent to the one shown in Fig. 5. (Fig. 5 can also be viewed as a system model during equalizer training.) Assume further that the transmitted signal, the NEXT, and the echo are uncorrelated with one another. Then, by examining Fig. 5, we can write down an expression for the mean-squared decision-point error. For this we first note that the combined impulse response of the channel and the DFE forward filter is, after T-spaced sampling, given by $M'HK\boldsymbol{h}_f$; where H is a lower-triangular Toeplitz matrix having $\boldsymbol{h}$ as its first column (padded with trailing zeros as needed), K is an L_f-column matrix composed of the $ik+1$st ($i = 0, 1, 2, \cdots, L_f - 1$) columns of identity matrix, and M a matrix composed of the $im+1$st ($i = 0, 1, 2, \cdots$) columns of the identity matrix. The matrix K accounts for the k:1 ratio between the forward filter tap spacing and the channel sampling period, and the matrix M the m:1 ratio between the symbol period and the channel sampling period. Similarly, the combined impulse response of the echo path and the DFE forward filter is $M'H_eK\boldsymbol{h}_f$ where H_e is a lower-triangular Toeplitz matrix having $\boldsymbol{h}_e$ as its first column (padded with trailing zeros as needed). The mean-squared decision-point error is thus given by

$$\begin{aligned} &E\{\epsilon_{iT}^2 | \boldsymbol{h}_f, \boldsymbol{h}_b, \boldsymbol{h}_g\} \\ &\quad = E\{[(M'HK\boldsymbol{h}_f - P_b\boldsymbol{h}_b - \boldsymbol{e}_c)'\boldsymbol{a}_{iT} \\ &\qquad + (M'H_eK\boldsymbol{h}_f - P_g\boldsymbol{h}_g)'\boldsymbol{b}_{iT} + \boldsymbol{h}_f'K'\boldsymbol{x}_{iT}]^2\} \\ &\quad = (M'HK\boldsymbol{h}_f - P_b\boldsymbol{h}_b - \boldsymbol{e}_c)'R_a(M'HK\boldsymbol{h}_f - P_b\boldsymbol{h}_b - \boldsymbol{e}_c) \\ &\qquad + (M'H_eK\boldsymbol{h}_f - P_g\boldsymbol{h}_g)'R_b(M'H_eK\boldsymbol{h}_f - P_g\boldsymbol{h}_g) \\ &\qquad + \boldsymbol{h}_f'K'R_xK\boldsymbol{h}_f; \end{aligned} \tag{1}$$

where E denotes the expectation operation; $'$ denotes matrix transpose; $\boldsymbol{x}_{iT}$, $\boldsymbol{a}_{iT}$, and $\boldsymbol{b}_{iT}$ are, respectively, vectors of noise samples and far-end and near-end signal symbols, arranged in reverse time order and led with x_{iT}, a_{iT}, and b_{iT}, respectively; R_x, R_a, and R_b are the corresponding autocorrelation matrices; $\boldsymbol{e}_c$ is the $c+1$st column of the identity matrix; P_b is a matrix composed of the $c+2$nd through the $c+L_b+1$st columns of the identity matrix (which accounts for the $(c+1)$-symbol delay associated with the DFE feedback filter); and P_g is a matrix composed of the first L_g columns of the identity matrix (to account for the length of the EC). Some of the vectors and matrices in (1) can be of infinite dimension.

The mean-squared error can also be expressed in terms of frequency-domain quantities, but the potential cyclostationarity of

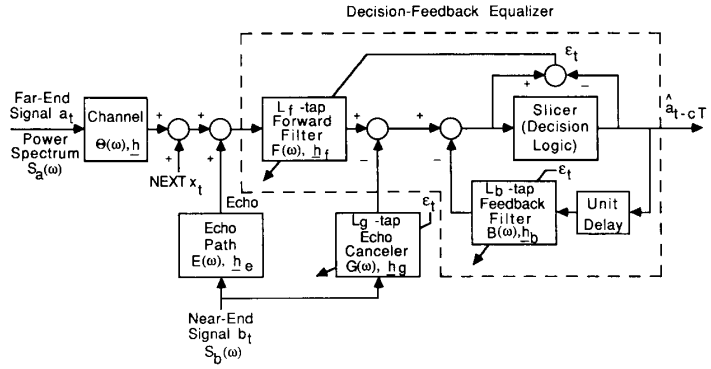

Fig. 4. A conceptualized DSL transmission path associated with the jointly adapted transceiver structure of Fig. 2.

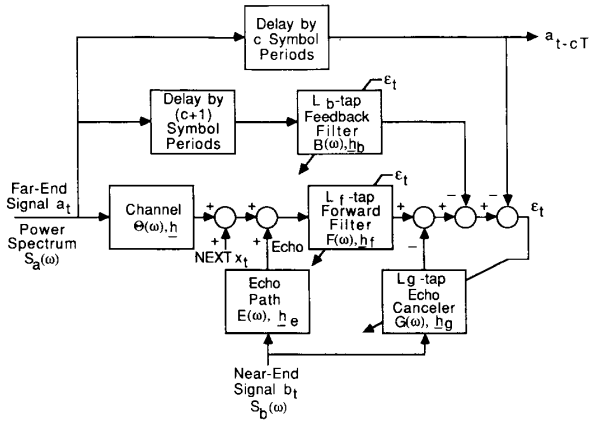

Fig. 5. A mathematically equivalent transmission path to that depicted in Fig. 4 for the jointly adapted transceiver structure.

the T/m-sampled self-NEXT sequence makes the formulation a little complicated. If we assume that this sequence is stationary, then

$$E\{\epsilon_{iT}^2 | F(\omega), B(\omega), G(\omega)\}$$

$$= \frac{T}{2\pi}\int_0^{2\pi/T}\left(S_a(\omega)\left|\frac{1}{m}\sum_{i=0}^{m-1}\Theta(\omega_i)F(\omega_i) - B(\omega)e^{-j\omega(c+1)T} - e^{-j\omega cT}\right|^2 + S_b(\omega)\left|\frac{1}{m}\sum_{i=0}^{m-1}E(\omega_i)F(\omega_i) - G(\omega)\right|^2 + \frac{1}{m}\sum_{i=0}^{m-1}S_x(\omega_i)|F(\omega_i)|^2\right)d\omega \quad (2)$$

where $\omega_i = \omega + 2\pi i/T$ and $S_x(\omega)$ is the NEXT power spectrum. This expression can be derived from (1) and Fig. 5 by Parseval's relation and the sampling theory [6]. We choose to work with the time-domain formulation because it is more convenient.

To minimize $E\{\epsilon_{iT}^2 | \boldsymbol{h}_f, \boldsymbol{h}_b, \boldsymbol{h}_g\}$, note that (1) can be rewritten simply as

$$E\{\epsilon_{iT}^2 | \boldsymbol{h}_f, \boldsymbol{h}_b, \boldsymbol{h}_g\} = \boldsymbol{\eta}' R \boldsymbol{\eta} - 2\boldsymbol{\rho}'\boldsymbol{\eta} + \sigma_a^2 \quad (3a)$$

where

$$\sigma_a^2 = \boldsymbol{e}_c' R_a \boldsymbol{e}_c \quad (3b)$$

(i.e., the diagonal element of R_a, or the far-end signal power);

$$\boldsymbol{\eta} = [\boldsymbol{h}_f' \vdots \boldsymbol{h}_b' \vdots \boldsymbol{h}_g']'; \quad (3c)$$

$$\boldsymbol{\rho}' = \boldsymbol{e}_c' R_a [M'HK \vdots -P_b \vdots 0]; \quad (3d)$$

and

$$R = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{12}' & R_{22} & 0 \\ R_{13}' & 0 & R_{33} \end{bmatrix} \quad (3e)$$

with

$$R_{11} = K'H'MR_aM'HK + K'H_e'MR_bM'H_eK + K'R_xK, \quad (3f)$$

$$R_{12} = -K'H'MR_aP_b, \quad R_{13} = -K'H_e'MR_bP_g, \quad (3g)$$

and

$$R_{22} = P_b'R_aP_b, \quad R_{33} = P_g'R_bP_g. \quad (3h)$$

Setting to zero the gradient of (3a) with respect to $\boldsymbol{\eta}$, we obtain the MMSE filters as

$$\boldsymbol{\eta} = R^{-1}\boldsymbol{\rho}. \quad (4)$$

The MMSE is then

$$\sigma_\epsilon^2 \equiv \min_{\boldsymbol{h}_g, \boldsymbol{h}_b, \boldsymbol{h}_f} E\{\epsilon_{iT}^2 | \boldsymbol{h}_f, \boldsymbol{h}_b, \boldsymbol{h}_g\} = \sigma_a^2 - \boldsymbol{\rho}'R^{-1}\boldsymbol{\rho}. \quad (5)$$

By the quadratic nature of $E\{\epsilon_{iT}^2\}$ as a function of $\boldsymbol{\eta}$, σ_ϵ^2 is a monotone nonincreasing function of the lengths of the EC and the DFE.

For numerical computation, it can be more manageable to have the above solution given in terms of the original parameters of (1). This can be done with fairly straightforward algebraic manipulations, which we outline in the Appendix. The result is as follows:

$$\sigma_\epsilon^2 = \sigma_a^2 - \boldsymbol{e}_c'R_aP_b(P_b'R_aP_b)^{-1}P_b'R_a\boldsymbol{e}_c - \boldsymbol{\rho}_1'R_1^{-1}\boldsymbol{\rho}_1, \quad (6a)$$

where

$$R_1 = K'R_xK + K'H'M[R_a - R_aP_b(P_b'R_aP_b)^{-1}P_b'R_a]M'HK + K'H_e'M[R_b - R_bP_g(P_g'R_bP_g)^{-1}P_g'R_b]M'H_eK \quad (6b)$$

and

$$\boldsymbol{\rho}_1 = K'H'M[R_a - R_aP_b(P_b'R_aP_b)^{-1}P_b'R_a]\boldsymbol{e}_c. \quad (6c)$$

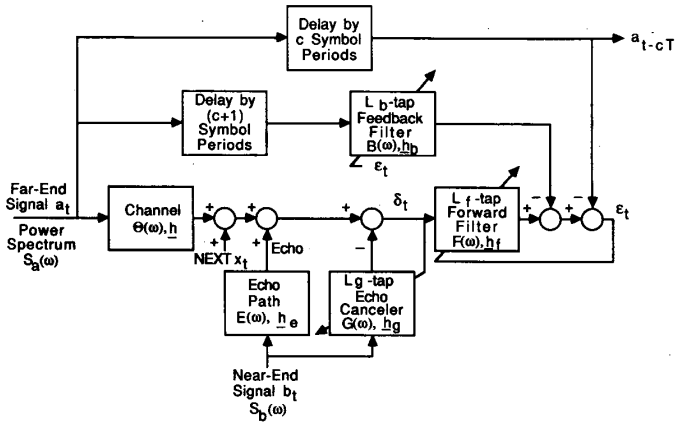

Fig. 6. A mathematically equivalent transmission path corresponding to the separately adapted transceiver structure.

Also, the optimal EC and DFE coefficients are given by

$$\boldsymbol{h}_f = R_1^{-1}\boldsymbol{\rho}_1, \tag{7a}$$

$$\boldsymbol{h}_b = (P_b'R_aP_b)^{-1}P_b'R_a(M'HK\boldsymbol{h}_f - \boldsymbol{e}_c), \tag{7b}$$

and

$$\boldsymbol{h}_g = (P_g'R_bP_g)^{-1}P_g'R_bM'H_eK\boldsymbol{h}_f. \tag{7c}$$

These equations may appear formidable at the first sight. A closer look reveals that there are several recurring quantities, namely, $M'HK$, $M'H_eK$, $R_aP_b(P_b'R_aP_b)^{-1}P_b'R_a$ or $(P_b'R_aP_b)^{-1}P_b'R_a$, and $R_bP_g(P_g'R_bP_g)^{-1}P_g'R_b$ or $(P_g'R_bP_g)^{-1}P_g'R_b$. The first and the second quantities denote the filtering functions of the channel and the echo path, respectively, as affected by the fractionally spaced sampling; and the third and the fourth reflect the effects of the DFE feedback filter and the echo canceller, respectively. Note that $M'HK$ and $M'H_eK$ are both block Toeplitz matrices with $(k \times m)$-sized blocks. The underlying structure in (6) and (7) may be easier to perceive by considering the case of binary symbol sequences, which we do in Section III.A.

B. The Separately Adapted Structure

Consider now the transceiver structure of Fig. 3. The corresponding "mathematically equivalent transmission path" is depicted in Fig. 6. Note the difference in the relative position between the EC and the DFE forward filter as compared to Fig. 5. Unlike in the jointly adapted structure, the EC is now assumed to have a fractional tap spacing T/m. While the transceiver structure is different, we shall be using substantially the same notations as for the previous one. This is for ease of cross reference as well as to reduce the notational burden. It should cause no confusion.

Let us first study the minimization of the mean-squared decision-point error ϵ by optimizing the DFE coefficients, assuming that the EC coefficients have been set. This mean-squared error is given by

$$\begin{aligned} E\{\epsilon_{iT}^2|\boldsymbol{h}_f, \boldsymbol{h}_b\} &= E\{[(M'HK\boldsymbol{h}_f - P_b\boldsymbol{h}_b - \boldsymbol{e}_c)'\boldsymbol{a}_{iT} \\ &\quad + \boldsymbol{h}_f'K'(H_e - H_g)'M\boldsymbol{b}_{iT} + \boldsymbol{h}_f'K'\boldsymbol{x}_{iT}]^2\} \\ &= (M'HK\boldsymbol{h}_f - P_b\boldsymbol{h}_b - \boldsymbol{e}_c)'R_a(M'HK\boldsymbol{h}_f \\ &\quad - P_b\boldsymbol{h}_b - \boldsymbol{e}_c) + \boldsymbol{h}_f'K'(H_e - H_g)'MR_bM' \\ &\quad \cdot (H_e - H_g)K\boldsymbol{h}_f + \boldsymbol{h}_f'KR_xK'\boldsymbol{h}_f \end{aligned} \tag{8}$$

where H_g is a lower triangular Toeplitz matrix having $P_g\boldsymbol{h}_g$ as its first column. Minimizing, we obtain the MMSE as

$$\sigma_\epsilon^2 = \sigma_a^2 - \boldsymbol{e}_c'R_aP_b(P_b'R_aP_b)^{-1}P_b'R_a\boldsymbol{e}_c - \boldsymbol{\rho}_2'R_2^{-1}\boldsymbol{\rho}_2, \tag{9a}$$

where

$$R_2 = K'R_xK + K'H'M[R_a - R_aP_b(P_b'R_aP_b)^{-1}P_b'R_a]M'HK + K'(H_e - H_g)'MR_bM'(H_e - H_g)K \tag{9b}$$

and

$$\boldsymbol{\rho}_2 = K'H'M[R_a - R_aP_b(P_b'R_aP_b)^{-1}P_b'R_a]\boldsymbol{e}_c. \tag{9c}$$

The optimal DFE is given by

$$\boldsymbol{h}_f = R_2^{-1}\boldsymbol{\rho}_2 \tag{10a}$$

and

$$\boldsymbol{h}_b = (P_b'R_aP_b)^{-1}P_b'R_a(M'HK\boldsymbol{h}_f - \boldsymbol{e}_c). \tag{10b}$$

Comparing (9) and (10) to (6) and (7), we see that the forms of the solutions are the same except for the effect of the echo canceller as manifested in the third terms of the expressions for R_1 and R_2.

Now we turn to the minimization of the EC output error δ. Note from Fig. 6 that this error contains three components, namely, the filtered far-end signal, the NEXT, and the residual echo, of which the first two are uncorrelated with the echo and are not affected by the EC. Thus the minimization of δ can be carried out by minimizing the residual echo alone. Now this residual echo sequence is not stationary, but cyclostationary with a period m due to the T/m sample spacing. A way to handle this is to minimize the sum of mean-squared errors of all the m phases of the error sequence. For this let M_j be the matrix obtained by prepending M with j rows of zeros. Then M_j' extracts the jth phase from an impulse response. For example, $M_j'\boldsymbol{h}_e$ gives the jth phase of the echo path response. Thus, we have the sum of mean-squared errors as

$$E\left\{\sum_{j=0}^{m-1}\delta_{iT+j/mT}^2|\boldsymbol{h}_g\right\} = E\left\{\sum_{j=0}^{m-1}[(\boldsymbol{h}_e - P_g\boldsymbol{h}_g)'M_j\boldsymbol{b}_{iT}]^2\right\}. \tag{11}$$

The MMSE EC can be solved by setting the derivative of (11) with respect to $\boldsymbol{h}_g$ to zero. It turns out that the solution is equivalent to that obtained by minimizing the mean-squared error for each of the m phases separately. Let L_j be the number of taps for the jth phase of the echo canceller. Then the optimal tap coefficients for this phase are given by

$$\boldsymbol{h}_j \equiv P_j'M_j'P_g\boldsymbol{h}_g = (P_j'R_bP_j)^{-1}P_j'R_bM_j'\boldsymbol{h}_e, \tag{12}$$

where P_j is a matrix composed of the first L_j columns of the identity matrix.

III. Application of MMSE Solutions to Various Line Codes

We consider four types of line codes, namely, binary-detected codes, precoded ternary partial-response codes, block codes yielding uncorrelated symbol sequences, and general block codes (of which partial-response codes can be viewed as special cases) which yield correlated symbol sequences. Examples of the first type of codes are the (*N*)RZ ([*n*on-]*r*eturn-to-*z*ero – coding 1 as +1 and 0 as −1; waveform returns or does not return to zero depending on the duty cycle) [8] and the nonprecoded partial-response codes [9]; that of the second type are the precoded dicode and the precoded MDB (*m*odified *d*uo*b*inary) [9]; that of the third are the 3B2T (3 *b*inary-to-2 *t*ernary, excluding 00) and the 2B1Q (2 *b*inary-to-1 *q*uarternary); and that of the fourth, the MS43 [10], [13], the MMS43 [11], and the DI43 [12]. All these codes have been considered for Basic Access ISDN DSL use. For the first type of codes, the "nonlinear operation" in a line coder (as shown in Fig. 1) is just a straight-through connection. For (*N*)RZ, the "partial-response filter" therein is also a straight-through connection. For the second type of codes, the nonlinear operation does the precoding and the partial-response filter does what the name implies. For the third and the fourth types of codes, the nonlinear operation does all the coding and the partial-response filter is a straight-through connection.

Again refer to Fig. 1. Taking the right-hand side as the near end, we have marked three points in the figure as A_1, A_2, and B, respectively. For all four types of codes, B corresponds to the point of near-end signal in Figs. 5 or 6 where the signal has a power spectrum S_b. For the second type of codes, A_2 corresponds to the point of far-end signal in Figs. 5 or 6 where the signal has a power spectrum S_a; and for the others, A_1. To further explain the correspondence between Fig. 1 and the mathematically equivalent transmission paths, consider the block labeled "channel" in Figs. 5 or 6. In terms of Fig. 1 components, the "channel" consists of, for the second type of codes, the far-end transmit filter, the far-end hybrid coupling, the line, the near-end hybrid coupling, and the near-end receive filter; and for the others, adding the far-end partial-response filter.

Let us make some comments on the self-NEXT noise. It undergoes a similar transmission mechanism as the far-end signal and the echo, in the sense that the source signals also pass through the same line coding, the same transmit filtering (both in the disturbing systems), and the same receive filtering (in the disturbed system). The only difference is that they are coupled into the disturbed system via crosstalk paths, which have different characteristics from the transmission and the echo paths. For nonsynchronized transmission systems, the crosstalk terms add noncoherently to form the total crosstalk noise. A common model for the power transfer function of NEXT coupling is $Kf^{3/2}$ where f denotes frequency and K is a parameter depending on the crosstalk strength [14]. A more exact characterization for some cases is given in [15]. If approximated by $Kf^{3/2}$, the one-crosstalker curve therein corresponds roughly to $K = 10^{-14}$ and the 49-crosstalker curve $K = 10^{-13}$, when f is given in Hertz.

A. Binary-Detected Codes and Block Codes Yielding Uncorrelated Symbol Sequences

For these line codes, $R_a = R_b = \sigma_a^2 I$. Computation of the MMSE solutions is relatively simple because $P_b' R_a P_b$, $P_g' R_b P_g$, and $P_j' R_b P_j$ are all diagonal matrices. Further, since $P_b' R_a \boldsymbol{e}_c = 0$, we have $\boldsymbol{\rho}_2 = \boldsymbol{\rho}_1 = \sigma_a^2 K' H' M \boldsymbol{e}_c$, which is a vector of kT/m-spaced channel impulse response samples arranged in reverse time order and led by the cm + 1st element of $\boldsymbol{h}$, the vector of T/m-spaced channel impulse response samples. In addition, the MMSE is given by $\sigma_\epsilon^2 = \sigma_a^2(1 - C)$ where $C = \boldsymbol{e}_c' M' H K \boldsymbol{h}_f$, for both transceiver structures.

For EC and DFE coefficients, consider first the jointly adapted structure. For the DFE feedback filter, we have $\boldsymbol{h}_b = P_b' M' H K \boldsymbol{h}_f$, i.e., we have come to the well-known result that the DFE feedback filter cancels out exactly the L_b taps of the forward-filtered (by $\boldsymbol{h}_f$) channel impulse response following the cursor. For the echo canceler, we have $\boldsymbol{h}_g = P_g' M' H_e K \boldsymbol{h}_f$, i.e., the EC cancels out exactly the first L_g samples of the forward-filtered echo path impulse response.

For the separately adapted structure, the DFE feedback filter again cancels out exactly the L_b taps of the filtered channel impulse response following the cursor. As to the echo canceller, we see from (12) that it cancels out exactly the first L_g samples of the original echo path impulse response.

B. Precoded Ternary Partial-Response Codes

For this type of codes we have $R_b = I$, because precoding is but an operation of cumulative modulo-2 summation which only makes a random input whiter and does not alter the spectral properties of an already white input [19], [20]. The code autocorrelation matrix R_a is nondiagonal but simple in structure.

Since $R_b = I$, the MMSE echo cancellers have the same properties as those for binary-detected codes. For $\sigma^2{}_\epsilon$ and the DFE, consider first the duobinary code which is characterized by a partial-response filter with coefficients (1, 1). R_a in this case is a tridiagonal Toeplitz matrix whose diagonal elements are all equal to 2 and whose first super- and first-subdiagonal elements are all equal to 1. For the computation of the MMSE solutions, we shall only comment on the structure of the quantities $P_b' R_a \boldsymbol{e}_c$ and $(P_b' R_a P_b)^{-1}$. It is not hard to see that

$$P_b' R_a \boldsymbol{e}_c = \boldsymbol{e}_0$$

where $\boldsymbol{e}_0$ denotes the first column of the identity matrix. The inverse of $P_b' R_a P_b$ can be obtained via triangularization as

$$(P_b' R_a P_b)^{-1} = D_1' D_2' \cdots D_{L_b-1}' \Sigma D_{L_b-1} \cdots D_2 D_1 \equiv D' \Sigma D$$

where $\Sigma = \text{diag}\,(1/2, 2/3, 3/4, \cdots, L_b/[L_b + 1])$ and D_i are elementary matrices having unity diagonals and a value $-i/(i+1)$ for the $(i+1)$,ith element. D is thus a lower triangular matrix with its ijth $(i \geq j)$ element equal to $(-1)^{i-j} j/i$, and $(P_b' R_a P_b)^{-1}$ a symmetric matrix with its ijth $(i \geq j)$ element equal to $(-1)^{i-j}(L_b - i + 1)j/(L_b + 1)$. This last observation readily enables us to unfold the structure of the quantity $R_a P_b (P_b' R_a P_b)^{-1} P_b' R_a$, which is not complicated but is somewhat cumbersome to describe—and is thus omitted here.

The above can be easily extended to interleaved duobinary codes, i.e., codes characterized by partial-response filters (1, 0, 1), (1, 0, 0, 1), etc. For these codes, we have a code autorrelation matrix equal to $I_{n+1} \otimes R_a^O$ where n denotes the number of zero coefficients in the partial-response filter associated with the interleaved code, I_{n+1} is the $(n+1) \times (n+1)$ identity matrix, $\otimes$ denotes matrix direct product [21] (to be further explained by an example below), and R_a^O is the original duobinary autocorrelation matrix described before. As a result, we obtain

$$P_b' R_a \boldsymbol{e}_c = \boldsymbol{e}_n$$

where $\boldsymbol{e}_n$ is the n + 1st column in the identity matrix; and

$$(P_b' R_a P_b)^{-1} = D_{(n)}' \Sigma_{(n)} D_{(n)}$$

where

$$D_{(n)} \sim I_{n-1} \otimes D \quad \text{and} \quad \Sigma_{(n)} \sim I_{n-1} \otimes \Sigma.$$

In the last few expressions, the subscript "(n)" signifies that the subscripted quantity is associated with an interleaved code having n zero coefficients in the partial-response filter, and the sign $\sim$ means that its lhs has the form described by its rhs. An example should help in clarifying the points. Consider the partial-response filter (1, 0, 1), for which $n = 1$. Consider the case where $L_b = 4$, which is an

integral multiple of $n+1$. Then we have

$$P_b'R_aP_b = \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix} = I_2 \otimes \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad (13)$$

$$D_{(1)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -\frac{1}{2} & 0 & 1 & 0 \\ 0 & -\frac{1}{2} & 0 & 1 \end{bmatrix} = I_2 \otimes \begin{bmatrix} 1 & 0 \\ -\frac{1}{2} & 1 \end{bmatrix}, \quad (14)$$

and

$$\Sigma_{(1)} = \text{diag}\left(\frac{1}{2}, \frac{1}{2}, \frac{2}{3}, \frac{2}{3}\right) = I_2 \otimes \text{diag}\left(\frac{1}{2}, \frac{2}{3}\right). \quad (15)$$

Now consider $L_b = 3$, which is not an integral multiple of $n+1$. It is not hard to see that the corresponding $P_b'R_aP_b$ can be obtained from (13) by deleting the first row and the first column. The corresponding $D_{(1)}$ and $\Sigma_{(1)}$ can be similarly obtained from (14) and (15), respectively. In general, the L_b-dimensional $D_{(n)}$ (and similarly for $P_b'R_aP_b$ and $\Sigma_{(n)}$) can be constructed by first forming the $\lceil L_b/(n+1) \rceil$-dimensional D corresponding to the duobinary code, taking its direct product with I_{n+1}, and then chopping off the first few rows and columns, if needed.

Now consider the bipolar code whose associated partial-response filter has coefficients $(1, -1)$. The code autocorrelation matrix is, as in the case of the duobinary code, tridiagonal Toeplitz with all the diagonal elements equal to 2. However, the first super- and first subdiagonal elements are now equal to -1. All the properties regarding the inverse of $P_b'R_aP_b$, including those pertaining to the interleaved codes, are similar to the corresponding ones for the duobinary code, except that the $(i+1)$, ith element of D_i is now equal to $i/(i+1)$, and hence the ijth element of D no longer has the $(-1)^{i-j}$ factor. The quantity $P_b'R_ae_c$ is easily seen to be equal to $-e_n$ where n is, again, the number of zero coefficients in the partial-response filter.

C. Block Codes Which Yield Correlated Symbol Sequences

These codes possess nonwhite spectra and nondiagonal autocorrelation matrices R_a and R_b. For them a brute-force implementation of (6) and (7), or (9), (10), and (12), to calculate the MMSE solution can be very computation-intensive. Happily, there are two techniques which we can call on to reduce the computation load. They both capitalize on the (block) Toeplitz structure of many of the constituent matrices in these equations. One is the Levinson–Trench algorithm [16]–[18] for efficient inversion of a Toeplitz matrix. A general method of inverting an $N \times N$ matrix, such as Gauss elimination followed by backward substitution, costs $O(N^3)$ multiplications. The Levinson–Trench algorithm does it in $O(N^2)$, which is an order of magnitude saving. The other technique is the FFT (fast Fourier transform [6]), useful in dealing with a multiplication with a large Toeplitz matrix. A multiplication with a large Toeplitz matrix is in effect a convolution operation, which can be done efficiently by multiplying the Fourier transforms of the two convolving sequences and inverse-transforming the result back into the time domain.

As an example, consider the $L_g \times L_g$ matrix R_1 defined in (6a). Let us discuss on how the quantities R_a, $(P_b'R_aP_b)^{-1}$, $R_aM'HK$ and $K'H'MR_aM'HK$ can be computed. The code autocorrelation matrix R_a can be obtained by inverse Fourier-transforming the code power spectrum S_a which can be calculated using an established method [13]. The inversion of the $L_b \times L_b$ Toeplitz matrix $P_b'R_aP_b$ can be accomplished by invoking the Levinson–Trench algorithm in case of a large L_b.

For the two other quantities, we first recall that $M'HK$ is a block Toeplitz matrix with blocksize $m \times k$. It can be viewed as representing a k-input m-output linear system. Before characterizing it in general, let us consider a simple example. In this example, let $k = 2$ and $m = 3$, i.e., we consider a $2T/3$-spaced forward filter. Then $M'HK$ can be viewed as representing a 2-input 3-output system as shown in Fig. 7. Note that the two input sequences are, respectively, the even and the odd subsequences of the original far-end signal sequence $\{a_i\}$. It is not hard to show that the autocorrelation matrix of the output multichannel sequence is exactly $K'H'MR_aM'HK$, a block Toeplitz matrix of 3×3 blocks. Define $\Theta_{2/3}(\omega)$ to be the 3×2 transfer function matrix of the 2-input 3-output system. Then the output has a matrix power spectrum as $\Theta_{2/3}(\omega)S_2^a(\omega)\Theta_{2/3}^*(\omega)$ where * denotes complex conjugate transpose and $S_2^a(\omega)$ is the matrix power spectrum of the 2-channel input process. The matrix $K'H'MR_aM'HK$ can therefore be formed from the inverse Fourier transform of $\Theta_{2/3}(\omega)S_2^a(\omega)\Theta_{2/3}^*(\omega)$ where $\Theta_{2/3}(\omega)$ can be calculated from the frequency response or the sampled impulse response of the channel, and $S_2^a(\omega)$ from the power spectrum or the autocorrelation sequence of the far-end signal. In a similar fashion, the quantity $R_aM'HK$, a Toeplitz matrix of 2×3 blocks, can be viewed as representing the cross-correlation between the 2-channel input and the 3-channel output, which has an associated matrix transfer function $S_2^a(\omega)\Theta_{2/3}^*(\omega)$. Generalizing to arbitrary k and m, we obtain a k-input m-output system as depicted in Fig. 8.

IV. Measurement of Transmission Performance

Let the decision thresholds be placed halfway between adjacent signal levels. The required SNR for detection in signal-independent additive Gaussian noise can be derived theoretically for the various line codes mentioned earlier for any target symbol error rate. Table I lists the results for the symbol error rates 10^{-6}, 10^{-7}, and 10^{-8}. Note that a symbol error may correspond to more than one bit error, in the case of block codes. However, for the block codes listed, the input–output mappings can be designed such that the average ratio between the number of erroneous bits and the number of erroneous symbols is less than two. For the range of error rates considered, a reduction of their values by a factor of two corresponds to an increase of the SNR requirement by less than 0.2 dB, and hence can be considered minor in a crude analysis.

In full-duplex DSL transmission, we see from Figs. 5 and 6 that the nominal SNR at the decision point after an MMSE echo cancellation and equalization is given by $\sigma_a^2/\sigma_\epsilon^2$. One may be easily tempted to tie this ratio hastily with the transmission error rate in the sense described in Table I. However, this is pertinent only when ϵ is Gaussian and independent of the signal.

Regarding the Gaussianness of ϵ, we note that it is natural to assume so in the absence of better and easily manageable noise models. In fact, after adequate echo cancellation and equalization, the residual echo and the residual ISI components of ϵ would both be a sum of many terms due, respectively, to the residual echo path and to the residual channel dispersion. Thus it should seem plausible that the residual echo and the residual ISI exhibit some kind of central limit behavior to justify this assumption. One may well add a certain amount of headroom to the SNR requirement listed in Table I to account for the potential non-Gaussianness of ϵ. However, this may be unnecessary as non-Gaussian noise can be more favorable than Gaussian (M. L. Honig and J. M. Cioffi, private communication). Note also that a common way of handling ISI (or any other noise) when it is non-Gaussian is to consider it as degrading the signal instead of adding to the noise. The worst degraded signal traces form an eye [7]. The eye can be computed mathematically or observed on an oscilloscope for a physical transmission system. With the Gaussianness assumption, such handling is not needed.

To see whether ϵ is independent of the signal, we have to examine two kinds of possible dependence. First, its conditional mean may be nonzero for some, and different for different, signal levels. (The nonzero conditional mean is a manifestation of the biased nature of linear MMSE estimates when the estimator input contains noise.) Second, its conditional variance may be different for different signal levels. To proceed, we note that σ_ϵ^2 can be decomposed as

$$\sigma_\epsilon^2 = E\{E[\epsilon_{iT}^2 | a_{(i-c)T} = \alpha]\}$$

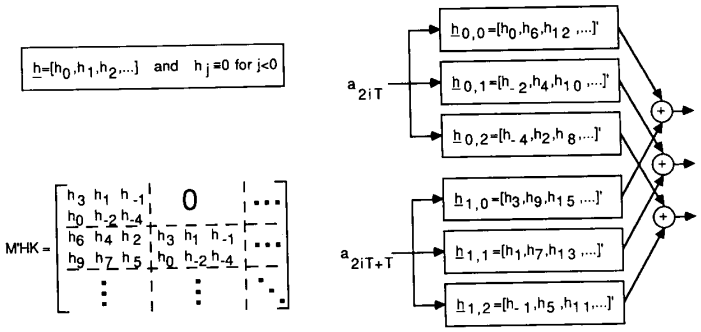

Fig. 7. Interpretation of $M'HK$ as representing a 2-input 3-output linear system when $k = 2$ and $m = 3$.

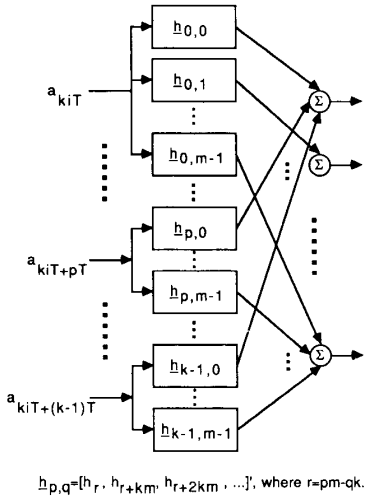

Fig. 8. Interpretation of $M'HK$ as representing a k-input m-output linear system for arbitrary k and m.

where the inner expectation is conditioned on $a_{(i-c)T} = \alpha$ and the outer expectation averages over all possible values of α. Now the inner expectation can be decomposed as

$$E[\epsilon_{iT}^2 | a_{(i-c)T} = \alpha] = [\mu(\alpha)]^2 + E\{[\epsilon_{iT} - \mu(\alpha)]^2 | a_{(i-c)T} = \alpha\}$$

where $\mu(\alpha) = E\{\epsilon_{iT} | a_{(i-c)T} = \alpha\}$, i.e., $\mu(\alpha)$ is the conditional bias in the estimated signal level. Hence, at the decision point, the true signal power is given by $E\{[\alpha + \mu(\alpha)]^2\}$ instead of σ_a^2, and the true noise power by $\sigma_\epsilon^2 - E\{[\mu(\alpha)]^2\}$ instead of σ_ϵ^2. To carry the derivation further, note that

$$\begin{aligned}\mu(\alpha) &= E\{(M'HK\boldsymbol{h}_f - P_b\boldsymbol{h}_b - \boldsymbol{e}_c)'\boldsymbol{a}_{iT} | a_{(i-c)T} = \alpha\} \\ &= (M'HK\boldsymbol{h}_f - P_b\boldsymbol{h}_b)'E\{\boldsymbol{a}_{iT} | a_{(i-c)T} = \alpha\} - \alpha. \quad (16)\end{aligned}$$

For codes whose successive symbols are conditionally uncorrelated for each signal level, we have $E\{\boldsymbol{a}_{iT} | a_{(i-c)T} = \alpha\} = \alpha\boldsymbol{e}_c$. Examples of such codes are the (N)RZ and the nonprecoded partial-response codes, the 3B2T, and the 2B1Q. For ternary codes whose statistics are symmetric about zero, we have $E\{\boldsymbol{a}_{iT} | a_{(i-c)T} = \alpha\} = \alpha R_a \boldsymbol{e}_c / \sigma_a^2$. Examples of such codes are the precoded dicode, the precoded MDB, the MS43, the MMS43, and the DI43. Substituting these into (16), we get the interesting result that, for all the codes considered,

$$\mu(\alpha) = \frac{\alpha}{\sigma_a^2}\boldsymbol{\rho}'\boldsymbol{\eta} - \alpha = -\alpha\sigma_\epsilon^2/\sigma_a^2,$$

for the jointly adapted structure; where we have used (3c), (3d), (4), and (5). For the separately adapted structure, the final expression for $\mu(\alpha)$ is the same. Consequently, for both transceiver structures, the true SNR is given by

$$\text{SNR} = \frac{\sigma_a^2}{\sigma_\epsilon^2} - 1, \quad (17)$$

which does not differ significantly from the nominal $\sigma_a^2/\sigma_\epsilon^2$ when we consider the 13.5 dB-and-up requirement listed in Table I. This shows that a signal-dependent mean of the error ϵ causes little concern.

On the other hand, a signal-dependent error variance can have more complicated implications. Their analysis also requires the development of more mathematical tools, especially in the case of block codes. Therefore, we elect to ignore them in this first study. Note that the binary-detected ([N]RZ and non-precoded partial-response) and the 2B1Q codes do not suffer from such a complication because the corresponding error variances are independent of signal levels due to the i.i.d. nature of the symbol sequences to be detected.

TABLE I
SYMBOL ERROR RATE VERSUS REQUIRED SNR (IN dB) FOR SOME LINE CODES IN TRANSMISSION OVER A GAUSSIAN NOISE CHANNEL

Line Code → / ↓ Error Rate	Binary-detected Codes [(N)RZ and Nonprecoded Partial-Response Codes]	Precoded Dicode, Precoded MDB	MS43, DI43	MMS43	3B2T	2B1Q
10^{-6}	13.5	16.8	17.9	17.9	18.5	20.8
10^{-7}	14.3	17.6	18.6	18.6	19.2	21.5
10^{-8}	15.0	18.2	19.3	19.2	19.9	22.2

V. An Ill-Behavior of the Separately Adapted Transceiver Structure

To illustrate the ill-behavior that a longer EC can result in a worse MMSE or SNR for the separately-adapted transceiver structure, consider a simple example where $R_a = R_b = I$. Let $k = m = 1$. Let there be no NEXT disturbance, i.e., $R_x = 0$. Let the channel impulse response be 1, 1, 0, 0, 0, $\cdots$, and the echo path impulse response be 0.5, 1, 0, 0, $\cdots$. Let the equalizer have a two-tap forward filter (i.e., $L_g = 2$) and no feedback filter (i.e., $L_b = 0$). And set the cursor delay c to 0. Then we have $\boldsymbol{\rho}_2' = [1\ 0]$. If there is no echo canceller, we also get

$$R_2 = H'H + (H_e - H_g)'(H_e - H_g) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} 1.25 & 0.5 \\ 0.5 & 1.25 \end{bmatrix};$$

and if there is a one-tap echo canceller,

$$R_2 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Thus for no echo canceller, we obtain $\sigma_\epsilon^2 = 1 - \boldsymbol{\rho}_2' R_2^{-1}\boldsymbol{\rho}_2 = 0.609$; while for a one-tap echo canceller, a worse MMSE as $\sigma_\epsilon^2 = 0.625$.

As can be seen, this peculiar behavior of the separately adapted transceiver structure arises because the EC coefficients affect the

correlation structure of the noise samples which the DFE downstream has to deal with. As a result, a longer EC, though reducing the power of the noise at its output, may alter the correlation structure of the noise to the disadvantage of the DFE. The jointly adapted structure does not suffer from this peculiarity because, as we have seen in Section II-A, there the mean-squared error is a quadratic function of the EC and the DFE coefficients and hence the MMSE is monotone nonincreasing with the lengths of the EC and the DFE. (The mean-squared decision-point error in the separately adapted structure is a quadratic function of the DFE coefficients but not of the EC coefficients.)

Some may suspect that, for "real" (in the sense described in this paper) DSL channels, the above ill-behavior of the separately adapted structure would show up only for very short echo cancellers. However, in our study, we have observed it for echo cancellers covering several tens of symbol periods.

The above example can also be used to demonstrate that, with the same EC and DFE lengths, the jointly adapted structure is not necessarily superior to the separately adapted in terms of the MMSE. For this, let the echo canceller have two taps. Then the separately adapted transceiver can fully cancel the echo, yielding an MMSE $\sigma_\epsilon^2 = 0.333$; whereas the jointly adapted transceiver can only cancel the echo partially and results in $\sigma_\epsilon^2 = 0.4$.

VI. Conclusion

We have presented a theory for the study of the optimal performance, in MMSE sense, of full-duplex DSL transceivers incorporating echo cancellers and decision-feedback equalizers. Two transceiver structures were considered, one with jointly adapted echo canceler and decision-feedback equalizer, and the other separately adapted. This theory augments previously published results by allowing a colored input signal and a fractionally spaced multitap MMSE DFE forward filter. We also discussed on how the MMSE is related to the ubiquitous SNR measure and the concerns associated with using it to gauge the DSL performance. For the separately adapted structure, we demonstrated an ill-behavior that a longer echo canceller can result in a worse MMSE or SNR.

Computational considerations of the theoretical results were also addressed. Specifically, we investigated the properties of and efficient ways of calculating the MMSE solutions for various types of line codes, including precoded partial-response codes and general block codes. A companion paper [22] reports some simulation results. A related paper [23] will deal with the important limiting situation of infinitely long DFE's by extending the results derived herein.

Appendix

We outline a procedure to derive expressions for the MMSE σ_ϵ^2 and the MMSE filters $\boldsymbol{h}_f$, $\boldsymbol{h}_b$, and $\boldsymbol{h}_g$, for the jointly adapted transceiver structure, in terms of the original parameters of (1) from (4) and (5).

Note first that the matrix R is block-diagonalizable by

$$L \equiv \begin{bmatrix} I & 0 & 0 \\ -R_{22}^{-1}R_{12}' & I & 0 \\ -R_{33}^{-1}R_{13}' & 0 & I \end{bmatrix}$$

in the sense that $D \equiv L'RL = \text{diag}\,(R_1, R_{22}, R_{33})$ where R_1 is as given in (6b) and R_{12}, R_{13}, R_{22}, and R_{33} are as defined in (3g) and (3h). Thus,

$$\boldsymbol{\eta} = R^{-1}\boldsymbol{\rho} = L(L'RL)^{-1}L'\boldsymbol{\rho} = LD^{-1}L'\boldsymbol{\rho} \tag{A1}$$

and

$$\sigma_\epsilon^2 = \sigma_a^2 - \boldsymbol{\rho}'R^{-1}\boldsymbol{\rho} = \sigma_a^2 - \boldsymbol{\rho}'LD^{-1}L'\boldsymbol{\rho}. \tag{A2}$$

From (3d), we have

$$\boldsymbol{\rho}'L = [\boldsymbol{\rho}_1' \vdots -\boldsymbol{e}_c'R_aP_b \vdots 0] \tag{A3}$$

where $\boldsymbol{\rho}_1$ is as given in (6c). The desired expressions for σ_ϵ^2 and the MMSE filters can then be obtained from (A2) and (A1) with the aid of (A3).

References

[1] D. G. Messerschmitt, "Echo cancellation in speech and data transmission," *IEEE J. Select. Areas Commun.*, vol. SAC-2, pp. 283–297, Mar. 1984.

[2] S. U. H. Qureshi, "Adaptive equalization," *Proc. IEEE*, vol. 73, pp. 1349–1387, Sept. 1985.

[3] K. H. Mueller, "Combining echo cancellation and decision feedback equalization," *Bell Syst. Tech. J.*, vol. 58, pp. 491–500, Feb. 1979.

[4] D. D. Falconer, "Adaptive reference echo cancellation," *IEEE Trans. Commun.*, vol. COM-30, pp. 2083–2094, Sept. 1982.

[5] O. Agazzi, D. A. Hodges, and D. G. Messerschmitt, "Large-scale integration of hybrid-method digital subscriber loops," *IEEE Trans. Commun.*, vol. COM-30, pp. 2095–2108, Sept. 1982.

[6] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1975.

[7] Members of Technical Staff, AT&T Bell Laboratories, *Transmission Systems for Communications*, AT&T Bell Laboratories, 1982, 5th ed.

[8] J. C. Bellamy, *Digital Telephony*. New York: Wiley, 1982.

[9] P. Kabal and S. Pasupathy, "Partial-response signaling," *IEEE Trans. Commun.*, vol. COM-23, pp. 921–934, Sept. 1975.

[10] P. A. Franaszek, "Sequence-state coding for digital transmission," *Bell Syst. Tech. J.*, vol. 47, pp. 143–157, Jan. 1968.

[11] ITT Telecom, Raleigh, North Carolina, USA, and Standard Elektrik Lorenz AG, Stuttgart, W. Germany, "Digital subscriber loop carrier network interface specification," ANSI/ECSA Telecommunications Working Group T1D1.3 Contribution 85-061, Apr. 1985.

[12] P. E. Fleischer and L. Wu, "Novel block codes for DSL applications," *IEEE Global Telecommun. Conf. Rec.*, 1985, pp. 1329–1334.

[13] G. L. Cariolaro and G. P. Tronca, "Spectra of block coded digital signals," *IEEE Trans. Commun.*, vol. COM-22, pp. 1555–1564, Oct. 1974.

[14] A. J. Gibbs and R. Addie, "The covariance of near end crosstalk and its application to PCM system engineering in multipair cable," *IEEE Trans. Commun.*, vol. COM-27, pp. 469–477, Feb. 1979.

[15] Bell Communications Research, Inc., "ISDN basic access digital subscriber lines," Tech. Ref. TR-TSY-000393, issue 1, May 1988.

[16] B. Friedlander, "Lattice filters for adaptive processing," *Proc. IEEE*, vol. 70, pp. 829–867, Aug. 1982.

[17] J. Makhoul, "Linear prediction: A tutorial review," *Proc. IEEE*, vol. 63, pp. 561–580, Apr. 1975.

[18] W. F. Trench, "An algorithm for the inversion of finite Toeplitz matrices," *SIAM J.*, vol. 12, pp. 515–522, Sept. 1964.

[19] D. G. Leeper, "A universal digital data scrambler," *Bell Syst. Tech. J.*, vol. 52, pp. 1851–1865, Dec. 1973.

[20] O. Agazzi, D. G. Messerschmitt, and D. A. Hodges, "Nonlinear echo cancellation of data signals," *IEEE Trans. Commun.*, vol. COM-30, pp. 2421–2433, Nov. 1982.

[21] W. K. Pratt, *Digital Image Processing*. New York: Wiley, 1978, sect. 5.1.

[22] D. W. Lin, "Minimum mean-squared error echo cancellation and equalization for digital subscriber line transmission: Part II—A simulation study," *IEEE Trans. Commun.*, vol. 38, pp. 39–45, Jan. 1990.

[23] ——, "Minimum mean-squared error decision-feedback equalization for digital subscriber line transmission with possibly correlated line codes," *IEEE Trans. Commun.*, to be published.

★

David W. Lin (S'78–M'82–SM'88) received the B.S. degree in 1975 from National Chiao Tung University, Hsinchu, Taiwan, China, and the M.S. and Ph.D. degrees in 1979 and 1981, respectively, from the University of Southern California, Los Angeles, all in electrical engineering.

He joined AT&T Bell Laboratories, Holmdel, NJ, in August 1981, and performed research in the areas of digital adaptive filtering and echo cancellation. Since January 1984, he has been with Bell Communications Research, Inc., and has worked on digital subscriber line transmission, speech coding, and video coding. His research interest includes various topics in signal processing and digital communication.