AT&T Technical Journal
Vol. 64, No. 1, January 1985
Printed in U.S.A.

Effects of Biases on Digitally Implemented Data-Driven Echo Cancelers

By J. M. CIOFFI* and J. J. WERNER†

(Manuscript received July 10, 1984)

In this paper the effects of biases, of a hardware origin, on the performance of a digitally implemented data-driven echo canceler are studied both analytically and experimentally. It is shown that, as a consequence of any such bias, the canceler tap weights can randomly drift; however, in contrast to voice-type cancelers and fractionally spaced equalizers, the data-driven canceler will not drift into instability. Nevertheless, the canceler's performance can be severely degraded, even for very small amounts of bias. The main result of this paper is a quantitative study of the canceler's performance as a function of the biases and the canceler's various design parameters, such as the number of tap coefficients and the step size used in the tap adjustment algorithm. Although the study concentrates on the biases introduced by two's-complement arithmetic, the results are general enough to be used with any type of arithmetic, provided that the biases introduced by these different types of arithmetic are known. Some of the analytical results have been verified experimentally, in real time, on a digital signal processor constructed at AT&T Bell Laboratories and AT&T Information Systems. Specifically, it is shown how the bias introduced by rounding the product of commercially available two's-complement multipliers can be eliminated by a proper choice of the values of the canceler's input symbols.

I. INTRODUCTION AND SUMMARY

The "tap-drifting" problem in fractionally spaced equalizers and voice-type echo cancelers is a manifestation of the presence of small

* AT&T Bell Laboratories; now at IBM Research. † AT&T Information Systems, Inc.

biases in the digital implementation of adaptive algorithms. These biases allow some of the equalizer and canceler tap coefficients to slowly grow in magnitude. As a result, the intermediate accumulated sums in the filtering computations also grow. Ultimately, either the tap coefficients or the intermediate-accumulated sum exceeds the boundaries of the digital representation ("overflow"), and the algorithms become unstable. The sensitivity of these algorithms to biases has been investigated, and several corrective actions against tap drifting have been proposed and successfully implemented.[1,2] Tap drifting can, in principle, be eliminated in its hardware origin by removing all the biases from the digital implementation. However, the elimination of biases is not always possible, especially if off-the-shelf devices are used in the design. For example, most commercially available arithmetic units use two's-complement arithmetic, which introduces biases in the computations when the digital words are reduced in length.

In this paper, we present a study of the tap-drifting problem in a data-driven canceler. This canceler is very attractive for two-wire full-duplex data transmission applications.* In its simplest configuration, for 4800-b/s full-duplex operation, the data-driven canceler requires only additions and subtractions for the computation of the filtering and updating algorithms. In this case no multiplier is required in the implementation of the canceler. However, a multiplier is required for higher speeds of transmission ($\geq$9600 b/s).

For most implementations, the effect of bias can be mitigated by using enough precision in the digital computations. However, such an approach will generally not be cost-effective, and in most practical digital implementations there will be some small biases due to the finite precision used in the computations. It is shown in this paper that, unlike voice-type cancelers,[4] data-driven cancelers cannot be driven to instability by these biases. Nevertheless, tap drifting always occurs to some extent and, after a sufficient period of time, introduces a degradation in the canceler's performance. This phenomenon has been studied both analytically and experimentally. We give a general formula that permits the degradation in the canceler's performance due to digital biases to be predicted. Real-time experimental results were obtained on a digital signal processor. The frequency components in the distortion introduced by the biases are shown to include spectral lines located at the origin and at multiples of the symbol rate. It is also shown that the performance of the biased canceler is degraded when the step size used in the updating of the tap coefficients de-

* The motivations for using echo cancellation in these applications are explained in Ref. 3. Section II of Ref. 3 also presents more details about the echo canceler structure used in the following study.

creases. This is in direct contrast to the behavior of the unbiased, infinite-precision canceler for which performance improves with decreasing step sizes. This phenomenon has also been observed in other applications.[5-7]

Special attention has been given to the degradation introduced by commercially available two's-complement multipliers. The naive use of such multipliers is shown to degrade the canceler's performance to unacceptable levels, even when large precision is used in the digital implementation. However, it is proved in this paper that proper choice of the values of the data symbols can completely eliminate the bias associated with the rounding of a two's-complement product. Two sets of binary symbols having this desired property are described.

The paper is organized as follows. In the next section the data-driven canceler used in the analysis and experiments is briefly described. In Section III we discuss the mathematical modeling of the biases, when two's-complement arithmetic is utilized in the canceler's updating algorithm. Quantitative results for the degradation in the canceler's performance are obtained in Section IV. The frequency-domain characteristics of the distortion introduced by the biases are studied in Section V. Finally, in Section VI we present some experimental results obtained on a digital signal processor, and we compare these results to the analytical results.

II. CANCELER DESCRIPTION

The data-driven passband echo canceler described in Refs. 8 and 3 synthesizes a signal of the form

$$s(t) = \text{Re}\left\{\sum_n A_n g(t - nT)e^{j\omega_c t}\right\}, \tag{1}$$

where $A_n = a_n + jb_n$ is the complex symbol to be transmitted, $g(t)$ is a (possibly complex) baseband signal, $\omega_c/2\pi$ is the carrier frequency, and Re denotes the real part of the quantity in brackets. It is shown in Ref. 8 that this signal can be generated by using the structure given in Fig. 1. The canceler consists of two transversal filters whose taps are spaced at intervals T', where $1/T'$ has to be at least twice the highest frequency in the signal $s(t)$ in (1). This condition makes the canceler Nyquist, that is, it can generate an exact replica of $s(t)$ at all frequencies.

After convergence, the tap coefficients are equal to the sampled values of the impulse responses of the in-phase and quadrature passband filters. For this reason the canceler is called an "in-band" canceler, distinguishing it from other possible structures that synthesize baseband equivalents of the channel. The symbol rotation at the input of the canceler ensures phase continuity of the carrier. For most

cases of practical interest, the relationship between ω_c and T is such that the rotated symbols a_n' and b_n' are similar to the symbols a_n and b_n. That is, if a_n and b_n are binary levels $\{\pm a\}$, then the rotated symbol levels will also be binary. Thus, the primes have been dropped in the ensuing analysis. The Digital-to-Analog (D/A) converter and low-pass filter at the output of the canceler perform the usual interpolation functions needed for further analog processing.

Since inputs are accepted by the canceler at a rate of $1/T'$, while the data symbols are only presented at a rate $1/T$, L − 1 zero symbols, where $L = T/T'$, are inserted between successive nonzero inputs to the canceler. Thus, only one of every L complex taps is active for each filter iteration, as shown in Fig. 1. The unnecessary computations associated with the zero symbols can be eliminated by grouping taps that act simultaneously into L parallel subcancelers, as seen in Fig. 2. Similarly, the echo channel can be considered as a parallel combination of L subchannels. Convergence of the canceler is achieved by minimizing the Mean-Squared Error (MSE) between the outputs of each subchannel and the corresponding subcanceler. The subcancelers are assumed to adapt independently, and this assumption was found to be in excellent agreement with experimental results. The MSE for the whole canceler is obtained by averaging MSEs of all the subcancelers. A more detailed analysis of the subcanceler structure is given in Ref. 8. The echo canceler's performance in the presence of channel impairments is studied in a companion paper.[3] Some of the definitions used in Ref. 3 that are needed in the sequel are briefly repeated. The error for the ith subcanceler at time $nT + iT'$ is

$$e_{n,i} = (\underline{r}_{1i} - \underline{c}_{n,i})^T \underline{a}_n + (\underline{r}_{2i} - \underline{d}_{n,i})^T \underline{b}_n + \xi_i, \tag{2}$$

where $i = 1, 2, \cdots, L$ and the superscript T denotes a transposed vector. In (2), $\underline{r}_{1i}$ and $\underline{r}_{2i}$ are the sampled in-phase and quadrature impulse response vectors of the ith subchannel; $\underline{c}_{n,i}$ and $\underline{d}_{n,i}$ are the in-phase and quadrature tap coefficient vectors of the ith subcanceler; and ξ_i is a signal that is uncorrelated with the data symbols (the far-end signal and noise). The following definitions are also needed:

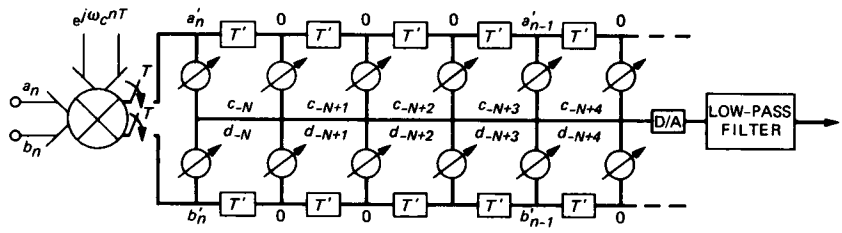
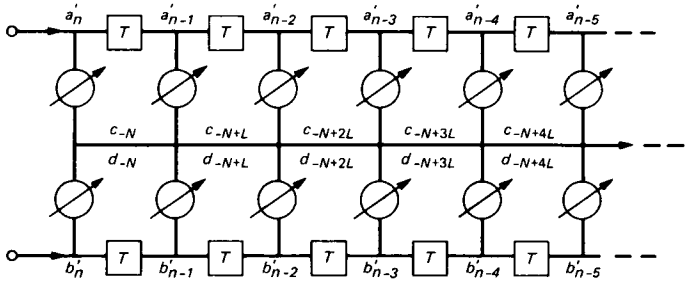
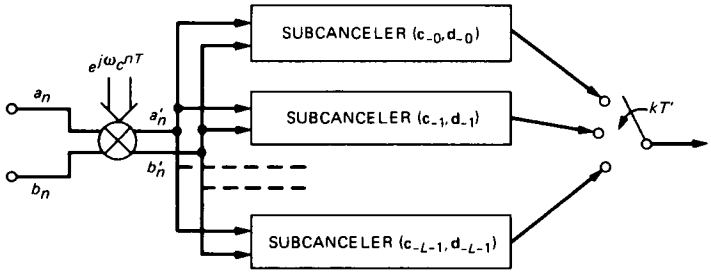


Fig. 1—In-band echo canceler structure.



(a)



(b)

Fig. 2—(a) Subcanceler structure. (b) Modified canceler structure.

$$\underline{A}_n = \underline{a}_n + j\underline{b}_n = \text{vector of complex input symbols,}$$

$$\underline{C}_{n,i} = \underline{c}_{n,i} + j\underline{d}_{n,i} = \text{vector of complex tap coefficients of the } i\text{th subcanceler.}$$

The above summary of the subcanceler structure is sufficient for the purpose of describing the effects of biases in this paper.

III. MATHEMATICAL MODEL OF BIAS

The stochastic-gradient algorithm for the adjustment of the ith subcanceler's complex tap coefficients is given by

$$\underline{C}_{n+1,i} = \underline{C}_{n,i} + \gamma e_{n,i}\underline{A}_n, \qquad i = 1, 2, \cdots, L, \tag{3}$$

where γ is the step size, and the error $e_{n,i}$ is given in (2). All of the entries (real and imaginary) in $\underline{A}_n$ are assumed to be discrete-valued symbols with zero mean. Biases are usually introduced in the evaluation of the correction term,

$$\underline{G}_{n,i} = \gamma e_{n,i}\underline{A}_n. \tag{4}$$

In the following, it is assumed that this expression is computed by using fractional two's-complement arithmetic, since this is the type of arithmetic most commonly used in digital signal processing. (Some of the properties of two's-complement arithmetic are discussed in Appendix B.) In a cost-conscious implementation, the quantity $\gamma e_{n,i}$ should be evaluated first, because it is the same for all the tap coefficients. First consideration is given to the case where (4) is computed without utilizing a multiplier, i.e., the scaled error, $\gamma e_{n,i}$, is obtained by using arithmetic right shifts since γ is always less than unity to ensure stability. Due to the finite precision of the digital representation, some of the lower bits of $e_{n,i}$ may be lost during the shifting operation, and consequently, a positive number will always decrease in magnitude and a negative number will always increase in magnitude. Therefore, negative bias is introduced, on the average, during the computation of $\gamma e_{n,i}$ and during "multiplication" by data symbols having value less than unity (since this operation also corresponds to right shifts). If all of the symbols are binary and chosen equal to ± 1, the updating algorithm in (3) is simply implemented by either adding or subtracting the scaled error to each tap coefficient. An explicit bias occurs during these operations if the quantity $\gamma e_{n,i}$ has "fallen out" of its register length. In this case a positive number becomes a true zero, but a negative number remains equal in magnitude to the Least-Significant Bit (LSB) of the digital representation, thus again introducing a negative mean bias. The correction factor in (4) can now be rewritten as

$$\underline{G}_{n,i} = (\gamma e_{n,i} + \Delta_{1n,i})\underline{A}_n + \underline{\Delta}_{2n,i} = \gamma e_{n,i}\underline{A}_n + \underline{B}_{n,i}, \qquad (5)$$

where $\Delta_{1n,i}$ is a random variable representing biases introduced in the evaluation of $\gamma e_{n,i}$; $\underline{\Delta}_{2n,i}$ is a complex random vector representing biases introduced in the computation of $(\gamma e_{n,i})\ \underline{A}_n$; and $\underline{B}_{n,i}$ is the total bias. The other quantities are assumed to be represented with infinite precision.

An expression similar to (5) is obtained when a two's-complement multiplier is used to compute the correction factor in the updating of the tap coefficients. When a negative product is truncated, the resulting number is always increased in magnitude. Conversely, when the product is positive, it is always decreased in magnitude. Therefore, truncation introduces a negative mean bias. Two's-complement rounding, on the other hand, always selects the number that is closest in magnitude to the exact product, independent of the product's sign. One exception, however, occurs when the double-precision product is equally close to two single-precision numbers. In this case the most positive number is always selected, independently of the product's

sign, and a positive bias is introduced in the computations.* This situation can occur quite frequently in data-driven echo cancelers using symbol values of $\pm 1/2$ or $\pm 1/4$. A remarkable property of data-driven cancelers is that the mean bias associated with this rounding can be eliminated by using the proper levels for the symbols. These levels must be chosen in a way such that the bias situation can never occur. Two sets of symbols which have the desired property are described in Appendix B.

The model represented by (5) also holds for other types of arithmetic. Furthermore, other effects such as biases in the Analog-to-Digital (A/D) conversion can also be accounted for by properly defining the random variables in $\Delta_{1n,i}$ and $\underline{\Delta}_{2n,i}$. In general, the statistics of these random variables will depend on γ, the statistics of $e_{n,i}$ and $\underline{A}_n$, and the type of digital implementation utilized. The characterization of the statistics of $\Delta_{1n,i}$ and $\underline{\Delta}_{2n,i}$ is a formidable problem that will not be addressed here. However, reasonable approximations permit the study of the canceler's performance degradation in the presence of digital biases. It will be shown later that both the mean and the variance of the random variables in (5) influence the canceler's performance. However, the effect of any nonzero-mean bias will, in general, be predominant. This quantity is discussed next.

From (5) the mean gradient estimate is given by

$$\langle \underline{G}_{n,i} \rangle = \langle \gamma e_{n,i} \underline{A}_n \rangle + \underline{B}_i, \tag{6}$$

where

$$\underline{B}_i = \langle \underline{B}_{n,i} \rangle = \langle \Delta_{1n,i} \underline{A}_n + \underline{\Delta}_{2n,i} \rangle. \tag{7}$$

The mean values of $\Delta_{1n,i}$ and $\underline{\Delta}_{2n,i}$ are generally not zero, as explained in the preceding discussion. In obtaining quantitative results, some assumptions concerning the mean values are made. First, all of the components of the vector $\langle \underline{\Delta}_{2n,i} \rangle$ corresponding to the ith subcanceler are assumed equal. This is a reasonable assumption since the same scaled error is used for the updating of all of the taps of a given subcanceler. The weighting of this term by $\underline{A}_n$ should yield the same average bias in steady-state operation. It is not assumed that the vector $\langle \underline{\Delta}_{2n,i} \rangle$ is the same for all the subcancelers, since the error, $e_{n,i}$, is a sample of a cyclostationary process whose statistics will depend upon the index i. As a consequence, the statistics of the product, $\gamma e_{n,i} A_n$, generally vary for different subcancelers.

* This type of bias is present in most of the commercially available two's-complement multipliers. However, the bias can be removed in a new design at the cost of a slight increase in the chip's complexity. This is achieved by first detecting the bias condition and then changing the rounding rules according to the sign of the product.

The influence of the term $\langle \Delta_{1n,i} \underline{A}_n \rangle$ in (7) is somewhat more difficult to assess. If the assumption is made that $\Delta_{1n,i}$ is uncorrelated with $\underline{A}_n$ and that $\langle \underline{A}_n \rangle$ is zero, then this quantity is also zero. However, there can be reasonably long sequences of symbols during which the mean value of $\underline{A}_n$ is not zero. Biases can accumulate during these sequences and introduce a degradation in the canceler's performance. We will not pursue this problem any further.

IV. EXCESS ERROR DUE TO BIAS

We will now show how biases can produce an increase in the MSE. It is assumed that no other impairments are present except, perhaps, for some uncorrelated noise added to the input signal. The mean tap coefficient fluctuations are investigated first. In the beginning of this section, the subscript i is dropped in the equations, but it will be understood that the analysis applied to a subcanceler. Combining (3) and (6), the mean tap vector evolves according to

$$\langle \underline{C}_{n+1} \rangle = \langle \underline{C}_n \rangle + \langle \gamma e_n \underline{A}_n \rangle + \underline{B}. \tag{8}$$

Using (2), we can write this expression as

$$\langle \underline{C}_{n+1} \rangle = (1 - \gamma A)\langle \underline{C}_n \rangle + (\underline{r}_1 + j\underline{r}_2)\gamma A + \underline{B}, \tag{9}$$

where $A = \langle a_n^2 \rangle = \langle b_n^2 \rangle$, and the $a_n's$ and $b_n's$ are assumed to be uncorrelated. The steady-state tap values are given by

$$\langle \underline{C}_\infty \rangle = \underline{r}_1 + j\underline{r}_2 + \frac{\underline{B}}{\gamma A}, \tag{10}$$

where the term $\underline{B}/\gamma A$ represents the tap deviation due to the mean bias $\underline{B}$. With $\Delta \underline{C}$ denoting the tap deviation from the optimum setting, one obtains

$$\Delta \underline{C} = \langle \underline{C}_\infty \rangle - \underline{C}_{\text{opt}} = \frac{\underline{B}}{\gamma A}. \tag{11}$$

Note that the bias is the same for each tap weight and that decreasing the step size γ will result in an increased mean tap deviation. This contrasts with the well-known results that decreasing the step size, for an infinite-precision canceler (without bias), will, in general, improve steady-state performance. The deviation is also proportional to the bias, which agrees with intuition. Recall that in a finite-precision environment the step size cannot be arbitrarily close to zero, and therefore the tap deviation in (11) cannot approach infinity. The only way to effectively have a zero step size is to stop updating, in which case the bias no longer affects the algorithms. When quantitative results are discussed in the following sections, the step sizes used in practice will usually be found to be several orders of magnitude larger

than the bias $\underline{B}$. Thus the tap offset in (11) is small. As a consequence, the tap coefficients will never overflow, and a data-driven canceler cannot become unstable. However, as will be shown later, a very small offset can severely degrade the MSE.

It is interesting to contrast the preceding results with those obtained for a voice-type canceler, which behaves similarly to fractionally spaced equalizers.[2] In this case it can be shown that the mean tap offset becomes

$$\Delta\underline{C} = \mathscr{A}^{-1}\frac{\underline{B}}{\gamma}, \tag{12}$$

where $\mathscr{A}$ is the input autocorrelation matrix of the data signal. Performing a spectral expansion of the mean tap deviation yields

$$\Delta\underline{C} = \frac{1}{\gamma}\sum_i \frac{1}{\gamma_i}\underline{f_i^T B f_i}, \tag{13}$$

where γ_i and $\underline{f_i}$ are the ith eigenvalue and eigenvector of $\mathscr{A}$, respectively. Small eigenvalues, corresponding to input frequency ranges of little or no energy, can make this term large, especially if $\underline{f_i^T B}$ is not small. Hence, for the voice-driven canceler, distortion due to biases can be expected to be concentrated in frequency ranges corresponding to little input energy. Furthermore, the tap offset in (13) can be much larger than the offset given in (11), so that some tap coefficients can overflow and make the canceler unstable.

The mean-squared error for the data-driven subcanceler is derived in Appendix A.* The expression for the ith subcanceler's steady-state MSE is given by

$$\langle e_{\infty,i}^2\rangle = \frac{\langle \xi_i^2\rangle + \dfrac{N}{\gamma}\langle \beta_i^2\rangle + \dfrac{2N}{\gamma^2 A}\langle \beta_i\rangle^2}{1-\gamma NA}, \tag{14}$$

where $\langle \xi_i^2\rangle$ is the minimum mean-squared error, N is the number of taps, and β_i is a random variable corresponding to one component (real or imaginary) of the vector $\underline{B}_{n,i}$ in (5). In the derivation of (14), it was assumed that $\langle \beta_i^2\rangle$ and $\langle \beta_i\rangle$ were constants and the same for all the entries of the vector $\underline{B}_{n,i}$. The MSE averaged over all L subcancelers is

* Another expression for the MSE (also derived in Appendix A) is obtained in the resonance case, when the step size has its optimum value for speed of convergence. This step size can only be used at start-up, when there is no double-talker, in which case convergence is so fast that the effects of biases are negligible. Therefore, we will not discuss this second case any further, and the MSE discussed in the sequel corresponds to the nonresonant case.

$$\langle e_\infty^2 \rangle = \frac{\langle \xi^2 \rangle + \dfrac{LN}{\gamma}\langle \beta^2 \rangle + \dfrac{2LN}{\gamma^2 A}\langle \beta^2 \rangle}{1 - \gamma NA}, \tag{15}$$

where we have defined

$$\langle \xi^2 \rangle = \frac{1}{L}\sum_{i=0}^{L-1} \langle \xi_i^2 \rangle, \qquad \langle \beta^2 \rangle = \frac{1}{L}\sum_{i=0}^{L-1} \langle \beta_i^2 \rangle, \tag{16}$$

and

$$\langle \beta \rangle^2 = \frac{1}{L}\sum_{i=1}^{L-1} \langle \beta_i \rangle^2.$$

Two quantities of interest can be derived from this expression. The first one is the *Bias-Performance Ratio* (BPR), which is defined as the ratio of the uncanceled echo's power to the MSE after cancellation, in the absence of a double-talker. This quantity is similar to the Echo-Return-Loss Enhancement (ERLE) studied in Ref. 3 and is particularly useful in comparing analytical and experimental results. The other quantity of interest is the signal-to-noise ratio (s/n), which is the ratio of the wanted signal power to the power of any additive noise after cancellation and which determines the receiver error rate. (From the receiver's viewpoint, the wanted signal is the double-talker, which acts as noise in the tap-adjustment algorithm.) From (15) the BPR is given by

$$\text{BPR} = \frac{P_e(1 - \gamma NA)}{\dfrac{LN}{\gamma}\langle \beta^2 \rangle + \dfrac{2LN}{\gamma^2 A}\langle \beta \rangle^2}, \tag{17}$$

where P_e is the power of the uncancelled echo. In deriving an expression for the s/n, it is assumed that the only component of the interfering signal is the wanted signal, i.e., there is no additive noise. The total signal power after cancellation is the MSE, (15), the power in the wanted signal is $P_s = \langle \xi^2 \rangle$, and the remaining power is noise, so that the s/n becomes

$$\text{s/n} = \frac{P_s}{\langle e_\infty^2 \rangle - P_s} = \frac{1 - \gamma NA}{\gamma NA + \dfrac{LN}{\gamma P_s}\langle \beta^2 \rangle + \dfrac{2LN}{\gamma^2 A P_s}\langle \beta \rangle^2}. \tag{18}$$

Under normal conditions of operation, the step size γ is very small, relative to $1/NA$, and the expressions in (17) and (18) can be approximated by

$$\text{BPR} \approx \frac{P_e A \gamma^2}{2LN\langle \beta \rangle^2} \tag{19}$$

and

$$\text{s/n} \approx \frac{P_s A \gamma^2}{2LN\langle\beta\rangle^2}, \tag{20}$$

where it has been assumed that $\langle\beta\rangle \neq 0$. The BPR and the s/n have the same expressions, except that the echo's power, P_e, in (19) is replaced by the wanted signal's power, P_s, in (20). Both quantities decrease with decreasing step size, which is in agreement with the preceding observations made about the mean tap behavior for small step size.

Equations (14) through (20) reflect degradation in the canceler's performance only under steady-state conditions. These steady-state quantities are not influenced by initial conditions. The complete equations for the MSE evolution with time are given by (55) and (64) in Appendix A. These equations arc strongly influenced, in their transient terms, by the canceler's initial state. An initial condition of particular interest is that which the condition that exists after the canceler has converged with no bias, and then a bias is introduced. As was discussed in Section III, this situation can arise as a result of certain nonrandom short-term statistics of the scrambled data sequence, A_n. Under these circumstances, the iterative MSE, given in (55), reduces to

$$\langle e_n^2\rangle = \frac{\langle\xi^2\rangle}{1-\gamma NA} + (1-[1-2\gamma A+2N\gamma^2A^2]^n)\frac{\dfrac{NL}{\gamma}\langle\beta^2\rangle + \dfrac{2NL}{\gamma^2 A}\langle\beta\rangle^2}{(1-\gamma NA)} + ([1-\gamma A]^n - [1-2\gamma A+2N\gamma^2A^2]^n)\frac{4N\langle\beta^2\rangle}{\gamma^2A(2N\gamma A-1)}, \tag{21}$$

and the evolution of the s/n is approximated by

$$s/n_n = \frac{P_s}{\langle e_n^2\rangle - P_s} \cong \frac{P_s\gamma^2 A}{P_s(\gamma NA) + [1-(1-2\gamma A)^n](2NL)\langle\beta\rangle^2}, \tag{22}$$

where it is assumed that the uncorrelated signal $\langle\xi^2\rangle$ consists only of the wanted signal with power, P_s. The approximation in (22) is obtained for a small step size. We will not pursue further the study of transient effects.

V. FREQUENCY ANALYSIS OF BIAS DISTORTION

For the data-driven canceler, distortion due to biases introduces spectral components concentrated in narrow frequency bands centered

around integer multiples of the symbol rate. To facilitate investigation of this phenomenon, the mean tap vector is again written in terms of its deviation from optimum

$$\langle \underline{C}_\infty \rangle = \underline{C}_{\text{opt},i} + \Delta \underline{C}_i, \tag{23}$$

where the subscript i is reintroduced to designate a subcanceler. The corresponding frequency response is the Fourier transform of the sequence obtained from the vector $\underline{C}_\infty$, and it is defined as

$$\langle C_{\infty,i}(\omega) \rangle = C_{\text{opt},i}(\omega) + \Delta C_i(\omega). \tag{24}$$

Since all components of the mean-bias vector $\underline{B}_i$ are assumed equal and

$$\Delta \underline{C}_i = \frac{\underline{B}_i}{\gamma A}, \tag{25}$$

then

$$\begin{aligned} \Delta C_i(\omega) &= \frac{B_i}{\gamma A} \sum_{n=0}^{N-1} e^{-j\omega nT}, \\ &= \frac{B_i}{\gamma A} e^{-j\omega(N-1)T/2} \frac{\sin[\omega NT/2]}{\sin[\omega T/2]}, \end{aligned} \tag{26}$$

where B_i is a component of $\underline{B}_i$ and $1/T$ is the symbol rate.

The amplitude of the distortion is the familiar periodic sinc function. From (26) it is seen that the distortion is concentrated at integer multiples of the symbol rate. As mentioned earlier, the bias vectors for each subcanceler need not be the same even if the bias components within any particular subcanceler are assumed constant. The corresponding tap-deviation spectrum can be expressed in terms of the tap-deviation spectra, $\Delta C_i(\omega)$, of each of the L subcancelers as

$$\Delta C(\omega) = \sum_{i=0}^{L-1} \Delta C_i(\omega) e^{-j\omega iT'}, \tag{27}$$

where $T' = T/L$.

If each $\Delta C_i(\omega)$ has the same shaping, but with a different magnitude, i.e.,

$$\Delta C_i(\omega) = k_i \Delta C_0(\omega), \qquad i = 0, \cdots, L-1, \tag{28}$$

where $k_0 = 1$ and the k_m are proportionality constants, then (27) becomes

$$\Delta C(\omega) = \Delta C_0(\omega) \sum_{i=0}^{L-1} k_i e^{-j\omega iT'}. \tag{29}$$

For the special case $k_m = 1$ for all m, the distortion becomes

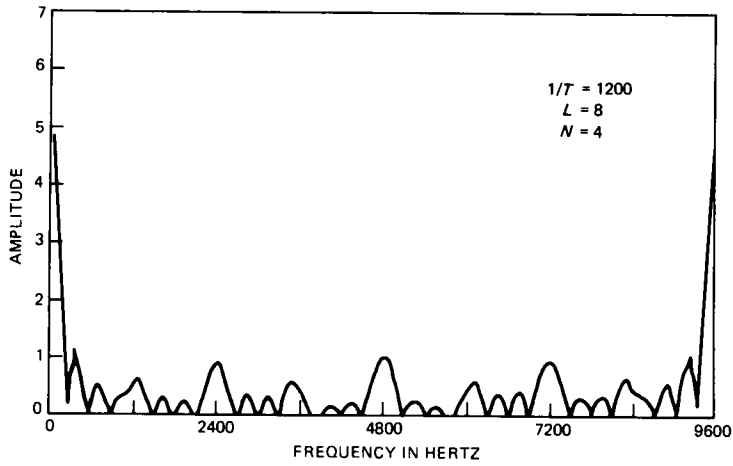


Fig. 3—Frequency distortion of mean tap coefficients.

concentrated at integer multiples of $1/T'$. Since all subcanceler biases are equal in this case, the structure becomes equivalent to one canceler at rate $1/T'$. As the value k_m varies, bias distortion is concentrated at integer multiples of the symbol rate, $1/T$.* One such example is shown in Fig. 3, where we have arbitrarily chosen values of k_m equal to 1, 0.5, 0.25, 0.5, 1, 0.5, 0.75, and 0.5. The study of many other examples shows that as a general rule of thumb, the spectral lines introduced by biases in a data-driven echo canceler composed of L subcancelers will be concentrated at integer multiples of the symbol rate. Experimental evidence verifying these findings will be presented in the next section.

VI. EXPERIMENTAL RESULTS

Several experiments were conducted on a digital signal processor, using a 12-bit two's-complement multiplier, to verify some of the qualitative and quantitative results obtained in the preceding sections. The echo canceler used in the experiments was operating at a sampling rate of 9600 samples per second, and a symbol rate of 1200 bauds. Thus, it could be implemented by using eight subcancelers. The validity of the expression for the BPR in (17) was verified by artificially inserting biases in the updating algorithm of an in-band echo canceler. A positive quantity equal to 2^{-11} was added periodically to the correction factor before updating the tap coefficients. Therefore, if this quantity was added every mT seconds, the equivalent mean bias per

* In certain rare cases, the biases can insert nulls at some particular multiples of the symbol rate.

symbol update was $(2^{-11})/m$. This bias was chosen large enough so that we could distinguish its effect from other digital effects such as round-off noise. The influence of the bias of the two's-complement multiplier was eliminated by using one of the sets of symbols described in Appendix B.

The measured values for the BPR are given in Fig. 4 for different values of m and as a function of the step size. The BPR has also been computed by using (17) and the corresponding curves are shown in Fig. 4. In Fig. 5, curves are also shown for the case in which the bias of the two's-complement multiplier was not eliminated by a proper choice of the data symbols. No artificial bias was added, and the computed curves were obtained by assuming a mean bias of 2^{-13}. (A

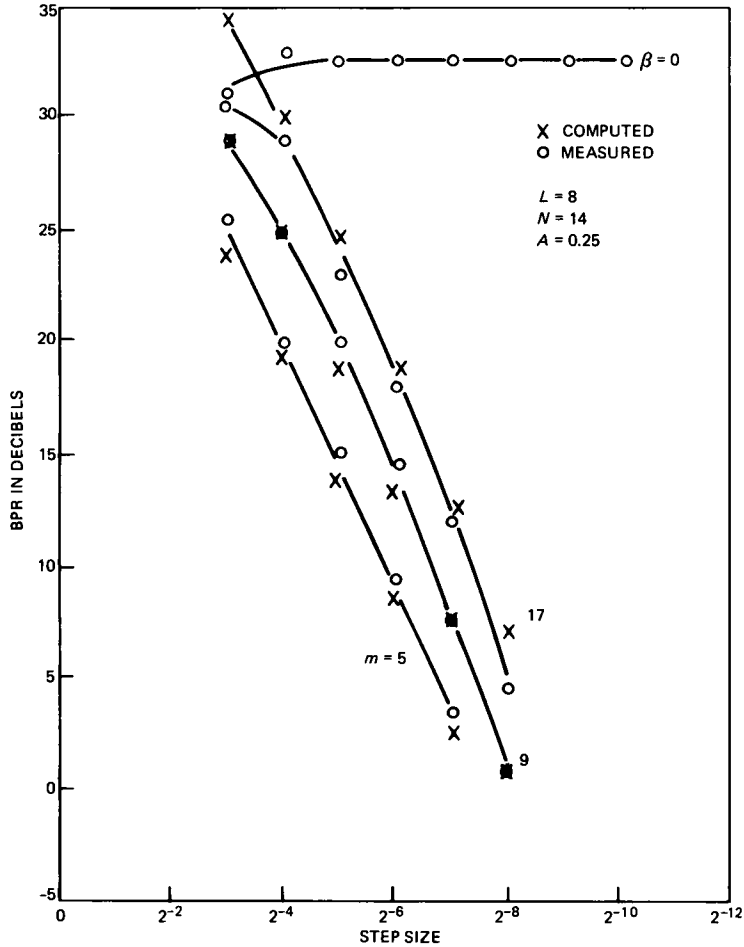


Fig. 4—Performance degradation due to simulated biases.

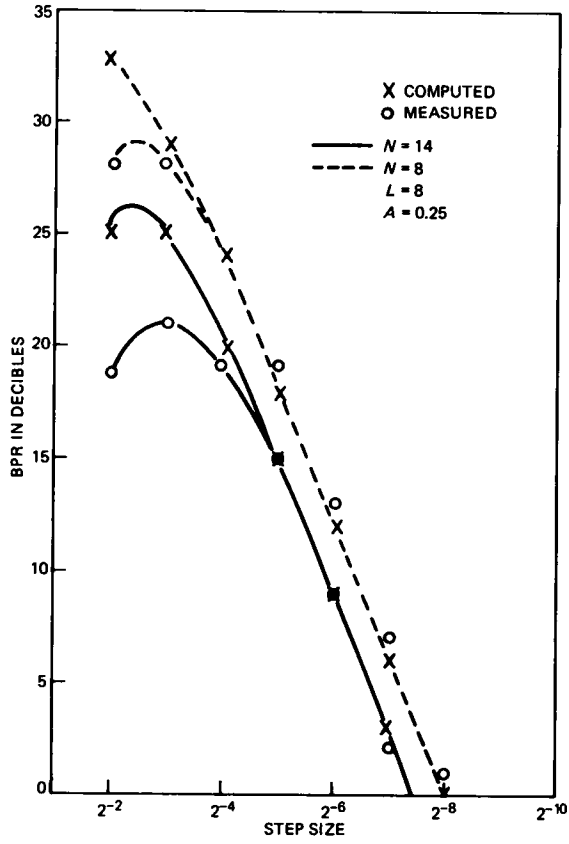


Fig. 5—Performance degradation due to rounding of a two's-complement multiplier.

bias of 2^{-12} was introduced whenever the bias situation occurred. With symbol values $\pm 1/2$, this situation was likely to occur half the time so that a mean bias of 2^{-13} was introduced in the algorithm.) For reasonably small step sizes, both the experimental and the theoretical curves decrease 6 dB for each factor-of-two decrease in step size. This is consistent with the expression for the BPR given in (19), and similar behavior can be expected for the s/n as shown in (20). Both these quantities go to zero when the step size goes to zero. This is in direct contrast to the behavior of a bias-free, infinite precision canceler for which performance improves with decreasing step sizes.

Although the two's-complement multiplier rounding bias was eliminated in the processor by using the sets of symbols described in Appendix B, there remained some other very small, unexplained biases. The effect of these biases on the BPR was negligible, as shown in Fig. 4. Nevertheless, they could be observed by studying the spec-

trum of the residual echo when no double-talker was present. This spectrum is the flat trace in Fig. 6, and the bell-shaped curve depicts the spectrum of the uncanceled echo. (Due to the small magnitude of the residual biases, it took several minutes of the canceler's operation to obtain the spectrum in Fig. 6. Immediately after convergence the peaks were very weak.) Notice that the peaks around the origin and at multiples of the symbol rate (1200 bauds in this case) produce exactly the kind of spectrum that was predicted by the analysis in Section V. These peaks would be much larger if the optimum sets of symbols described in Appendix B were not utilized. In Fig. 6, we also show the spectrum of the residual echo of a voice-type canceler that was implemented on the processor. Notice that for the voice-canceler spectrum, the residual echo's energy accumulates in the frequency regions where the uncanceled echo has little energy. This energy increases with time and ultimately, without some sort of compensation, the canceler diverges after several minutes of operation. Finally, for

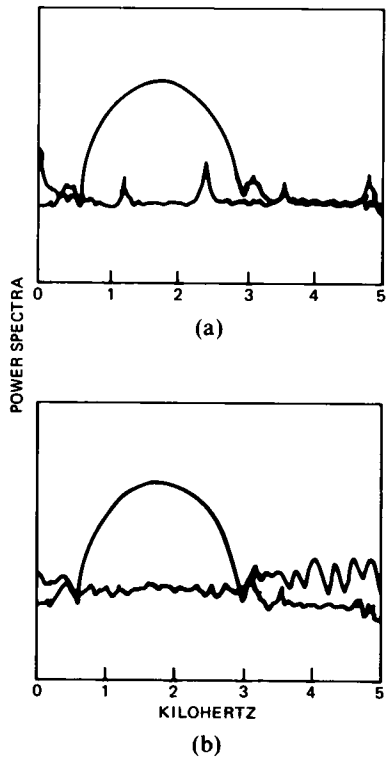


Fig. 6—Power spectrum of the residual error. (a) Data-driven echo canceler. (b) Voice-type echo canceler.

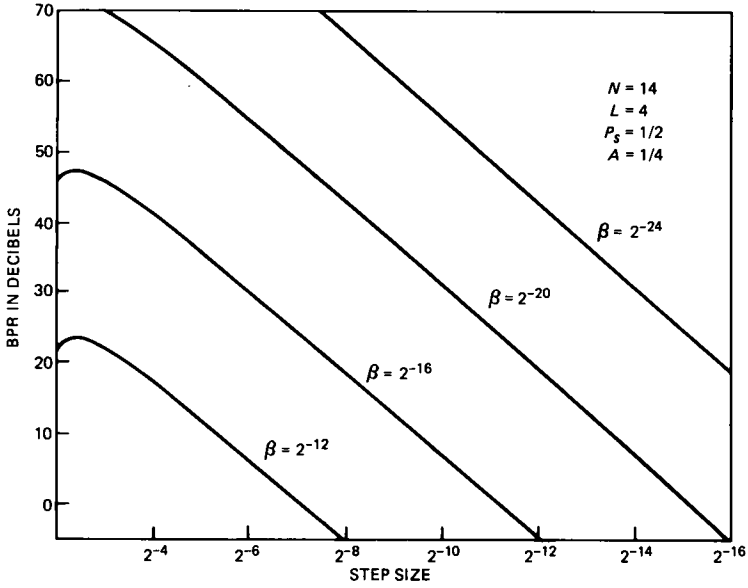


Fig. 7—Bias-performance ratio for various amounts of biases.

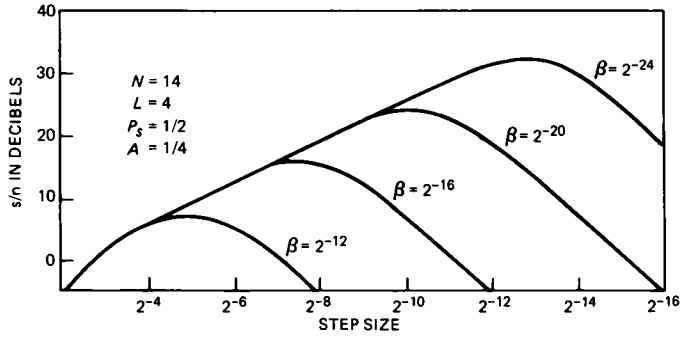


Fig. 8—Signal-to-noise degradation due to biases.

completeness, computed curves giving the BPR and the s/n for various values of biases are given in Figs. 7 and 8.

REFERENCES

1. J. J. Werner, "Control of Coefficient Drift for Fractionally Spaced Equalizers," US Patent 4,384,355, May 17, 1983.
2. R. D. Gitlin, H. C. Meadors, and S. B. Weinstein, "The Tap Leakage Algorithm: An Algorithm for the Stable Operation of a Digitally Implemented Fractionally Spaced Adaptive Equalizer," B.S.T.J., *61*, No. 8 (October 1982), pp. 1817–39.
3. J. J. Werner, "Effects of Channel Impairments on the Performance of an In-band Data-Driven Echo Canceler," AT&T Tech. J., this issue.
4. A. Weiss and D. Mitra, "Digital Adaptive Filters: Conditions for Convergence, Rates

of Convergence, Effects of Noise and Errors Arising from the Implementation," IEEE Trans. Inform. Theory, *IT-25*, No. 6 (November 1979), pp. 637–52.
5. R. D. Gitlin, J. E. Mazo, and M. G. Taylor, "On the Design of Gradient Algorithms for Digitally Implemented Adaptive Filters," IEEE Trans. Circuit Theory, *CT-20*, No. 2 (March 1973).
6. R. D. Gitlin and S. B. Weinstein, "On the Required Tap-Weight Precision for Digitally Implemented, Adaptive, Mean-Squared Equalizers," B.S.T.J., *58*, No. 2 (February 1979), pp. 301–21.
7. C. Caraiscos and B. Liu, "A Roundoff Error Analysis of the LMS Adaptive Algorithm," IEEE Trans. ASSP, *ASSP-32*, No. 2 (February 1984), pp. 34–41.
8. J. J. Werner, "An Echo-Cancellation-Based 4800 bps Full-Duplex DDD Modem," IEEE J. Selected Areas Commun., *SAC-2*, No. 5 (September 1984).

APPENDIX A

MSE Evolution

The following is an analysis of the evolution of the MSE for a data-driven echo canceler in the presence of digital biases. It is assumed that the input sequences $\underline{a}_n$ and $\underline{b}_n$ are white and uncorrelated with the bias, $\underline{\beta}_n$, and each other. It is also assumed that the canceler spans the entire length of the echo impulse response. The analysis applies to a subcanceler, but for simplicity of notation, the index i will be deleted.

The following definitions are used in the derivations:

$$\underline{B}_n = \underline{\beta}_n + j\underline{\beta}_n = \text{complex bias vector}^*$$

$$\underline{C}_n = \underline{c}_n + j\underline{d}_n = \text{complex tap vector}$$

$$\underline{A}_n = \underline{a}_n + j\underline{b}_n = \text{complex data vector}$$

$$\gamma = \text{step size in the adjustment algorithm}$$

$$\underline{r}_1 = \text{sampled in-phase channel impulse response vector} \tag{30}$$

$$\underline{r}_2 = \text{sampled quadrature channel impulse response vector} \tag{31}$$

$$\xi = \text{sampled uncorrelated interfering signal.} \tag{32}$$

The error at the nth iteration is given by

$$e_n = [\underline{r}_1 - \underline{c}_n]^T \underline{a}_n + [\underline{r}_2 - \underline{d}_n]^T b_n + \xi. \tag{33}$$

The in-phase and quadrature tap error vectors are defined by

$$\underline{\in}_{1,n} \triangleq \underline{r}_1 - \underline{c}_n \tag{34}$$

$$\underline{\in}_{2,n} \triangleq \underline{r}_2 - \underline{d}_n \tag{35}$$

* For ease of notation the real and imaginary parts of $\underline{B}_n$ are taken to be equal. Although this assumption is not strictly true, it will not modify the end result of the analysis.

so that (33) can be rewritten as

$$e_n = \underline{\in}_{1,n}^T \underline{a}_n + \underline{\in}_{2,n}^T \underline{b}_n + \xi, \tag{36}$$

and the MSE becomes

$$\langle e_{n+1}^2 \rangle = \langle \underline{\in}_{1,n+1}^T \underline{a}_{n+1} \underline{a}_{n+1}^T \underline{\in}_{1,n+1} \rangle + \langle \underline{\in}_{2,n+1}^T \underline{b}_{n+1} \underline{b}_{n+1}^T \underline{\in}_{2,n+1} \rangle + \langle \xi^2 \rangle. \tag{37}$$

Letting

$$A \triangleq \langle a_n^2 \rangle = \langle b_n^2 \rangle \tag{38}$$

and making the usual assumption that successive data vectors are uncorrelated, one obtains

$$\langle e_{n+1}^2 \rangle = A \langle \in_{1,n+1}^T \underline{\in}_{1,n+1} \rangle + A \langle \in_{2,n+1}^T \underline{\in}_{2,n+1} \rangle + \langle \xi^2 \rangle. \tag{39}$$

The biased stochastic-gradient algorithm can be expressed as

$$\underline{c}_{n+1} = \underline{c}_n + \gamma e_n \underline{a}_n + \underline{\beta}_n \tag{40}$$

$$\underline{d}_{n+1} = \underline{d}_n + \gamma e_n \underline{b}_n + \underline{\beta}_n . \tag{41}$$

Subtracting both sides of (40) and (41) from $\underline{r}_1$ and $\underline{r}_2$ yields

$$\underline{\in}_{1,n+1} = \underline{\in}_{1,n} - \gamma e_n \underline{a}_n - \underline{\beta}_n \tag{42}$$

$$\underline{\in}_{2,n+1} = \underline{\in}_{2,n} - \gamma e_n \underline{b}_n - \underline{\beta}_n . \tag{43}$$

Using (42) and (43) in (39) gives

$$\langle e_{n+1}^2 \rangle = A \langle (\underline{\in}_{1,n} - \gamma e_n \underline{a}_n - \underline{\beta}_n)^T (\underline{\in}_{1,n} - \gamma e_n \underline{a}_n - \underline{\beta}_n) \rangle + A \langle \underline{\in}_{2,n} - \gamma e_n \underline{b}_n - \underline{\beta}_n)^T (\underline{\in}_{2,n} - \gamma e_n \underline{b}_n - \underline{\beta}_n) \rangle + \langle \xi^2 \rangle . \tag{44}$$

After some algebra, and using (36), (38), and $\underline{a}_n^T \underline{a}_n = \underline{b}_n^T \underline{b}_n = NA$, we have

$$\langle e_{n+1}^2 \rangle = \langle e_n^2 \rangle [1 - 2\gamma A + 2\gamma^2 N A^2] + 2\gamma A \langle \xi^2 \rangle + 2AN \langle \beta_n^2 \rangle - 2A \langle \underline{\beta}_n^T \underline{\in}_{1,n} \rangle - 2A \langle \underline{\beta}_n^T \underline{\in}_{2,n} \rangle. \tag{45}$$

To proceed further, $\langle \underline{\beta}_n^T \underline{\in}_{1,n} \rangle$ and $\langle \underline{\beta}_n^T \underline{\in}_{2,n} \rangle$ must be evaluated. We will assume that $\underline{\beta}_n$ is uncorrelated with $\underline{\in}_{1,n}$ and $\underline{\in}_{2,n}$.

Using (42) and (36), and assuming again that the current tap error vector and the current data vector are independent, we have

$$\langle \underline{\beta} \rangle^T \langle \underline{\in}_{1,n+1} \rangle = (1 - \gamma A) \langle \underline{\beta} \rangle^T \langle \underline{\in}_{1,n} \rangle - N \langle \beta \rangle^2, \tag{46}$$

where it is assumed that the mean vector $\langle \underline{\beta}_n \rangle$ is a constant vector whose entries are all equal to $\langle \beta \rangle$. Solving the iteration (46) yields

$$\langle \underline{\beta} \rangle^T \langle \underline{\in}_{1,n} \rangle = (1 - \gamma A)^n \langle \underline{\beta} \rangle^T \underline{r}_1 - \frac{1 - (1 - \gamma A)^n}{\gamma A} N \langle \beta \rangle^2. \tag{47}$$

In (50) it is assumed that $\langle \underline{\in}_{1,0} \rangle = \underline{r}_1$, that is, the canceler is originally started with all of its tap coefficients equal to zero. A similar expression holds for the quadrature component.

The difference equation for the MSE in (45) now becomes

$$\langle e_{n+1}^2 \rangle = \langle e_n^2 \rangle [1 - 2\gamma A + 2N\gamma^2 A^2] + 2\gamma A \langle \xi^2 \rangle + 2AN\langle \beta_n^2 \rangle - 2A \left[(1 - \gamma A)^n \langle \underline{\beta}^T \rangle \underline{r}_1 - \frac{1 - (1 - \gamma A)^n}{\gamma A} N \langle \beta \rangle^2 \right] - 2A \left[(1 - \gamma A)^n \langle \underline{\beta}^T \rangle \underline{r}_2 - \frac{1 - (1 - \gamma A)^n}{\gamma A} N \langle \beta \rangle^2 \right]. \quad (48)$$

The above equation, although notationally complex, is only a first-order difference equation of the form

$$\langle e_{n+1}^2 \rangle = \langle e_n^2 \rangle \lambda_1 + k_1 - k_2 \lambda_2^n, \quad (49)$$

where

$$\lambda_1 = 1 - 2\gamma A + 2N\gamma^2 A^2 \quad (50)$$

$$\lambda_2 = 1 - \gamma A \quad (51)$$

$$k_1 = 2\gamma A \langle \xi^2 \rangle + 2AN\langle \beta^2 \rangle + \frac{4}{\gamma} N \langle \beta \rangle^2 \quad (52)$$

$$k_2 = 2A \langle \underline{\beta}^T \rangle \underline{r}_1 + 2A \langle \underline{\beta}^T \rangle \underline{r}_2 + \frac{4}{\gamma} N \langle \beta \rangle^2, \quad (53)$$

and where it is assumed that $\langle \beta_n^2 \rangle$ is a constant and equal to $\langle \beta^2 \rangle$. Assuming stability,

$$|\lambda_1| < 1 \quad \text{and} \quad |\lambda_2| < 1, \quad (54)$$

there are two cases to be considered in solving (53).

Case 1: $\lambda_1 \neq \lambda_2$ (nonresonant).

Assume

$$\langle e_n^2 \rangle = C_1 + C_2 \lambda_1^n + C_3 \lambda_2^n. \quad (55)$$

Then solving (49) by substitution, and assuming the initial condition $\langle e_0^2 \rangle$ gives

$$C_1 = \frac{k_1}{1 - \lambda_1} \quad (56)$$

$$C_3 = \frac{-k_2}{\lambda_2 - \lambda_1} = \frac{k_2}{\lambda_1 - \lambda_2} \quad (57)$$

$$C_2 = \langle e_0^2 \rangle - C_1 - C_3. \quad (58)$$

In terms of the original parameters in (48),

$$C_1 = \frac{\langle \xi^2 \rangle + \frac{N}{\gamma} \langle \beta^2 \rangle + \frac{2N}{\gamma^2 A} \langle \beta \rangle^2}{(1 - \gamma NA)} \tag{59}$$

$$C_3 = \frac{\frac{2}{\gamma} [\langle \underline{\beta}_T \rangle \underline{r}_1 + \langle \underline{\beta}^T \rangle \underline{r}_2] + \frac{4N}{\gamma^2 A} N \langle \beta \rangle^2}{(2N\gamma A - 1)} \tag{60}$$

$$C_2 = \langle e_0^2 \rangle - C_1 - C_3. \tag{61}$$

All of the parameters appearing in the expression for the MSE in (55) are now known. In the steady state, as n goes to infinity the MSE becomes

$$\langle e_\infty^2 \rangle = \frac{\langle \xi^2 \rangle + \frac{N}{\gamma} \langle \beta^2 \rangle + \frac{2N}{\gamma^2 A} \langle \beta \rangle^2}{1 - \gamma NA}. \tag{62}$$

Case 2: Resonance ($\gamma_1 = \gamma_2$).
In this case, from (50) and (51)

$$\gamma = \frac{1}{2NA}. \tag{63}$$

Assume

$$\langle e_n^2 \rangle = C_1 + C_2 \gamma_1^n + n C_e \gamma_1^n. \tag{64}$$

Solving (49) again by substitution, we get, after some algebra,

$$\begin{aligned} \langle e_n^2 \rangle &= 2\langle \xi^2 \rangle + 4NA \langle \beta^2 \rangle + 16N^2 A \langle \beta \rangle^2 \\ &\quad + [\langle e_0^2 \rangle - 2\langle \xi^2 \rangle - 4NA \langle \beta^2 \rangle - 16N^2 A \langle \beta \rangle^2] \left(1 - \frac{1}{2N}\right)^n \\ &\quad - n[2A[\langle \underline{\beta}^T \rangle \underline{r}_1 + \langle \underline{\beta}^T \rangle \underline{r}_2] + 8N^2 A \langle \beta \rangle^2] \left(1 - \frac{1}{2N}\right)^{n-1}. \end{aligned} \tag{65}$$

In this case the steady-state MSE is

$$\langle e_\infty^2 \rangle = 2\langle \xi^2 \rangle + 4NA \langle \beta^2 \rangle + 16N^2 A \langle \beta \rangle^2. \tag{66}$$

APPENDIX B

Bias-Free Two's-Complement Multiplication for a Data-Driven Canceler

In this Appendix, it is shown that the bias that occurs when rounding the product of a two's-complement multiplier can be eliminated by properly choosing the values of the binary input symbols. Let α be an (l+1)-bit fractional two's-complement number represented as

$$\alpha = a_0 a_1 a_2 \cdots a_l, \tag{67}$$

where $a_i = 0$ or 1, and a_l is the LSB of the number.
The numerical value of α is given by*

$$\alpha = -a_0 + \sum_{i=1}^{l} a_i 2^{-i}, \tag{68}$$

and the product of two such numbers becomes

$$\begin{aligned}\alpha \cdot x &= \left(-a_0 + \sum_{i=1}^{l} a_i 2^{-i}\right)\left(-x_0 + \sum_{k=1}^{l} x_k 2^{-k}\right) \\ &= a_0 x_0 - a_0 \sum_{k=1}^{l} x_k 2^{-k} - x_0 \sum_{i=1}^{l} a_i 2^{-i} + \sum_{i=1}^{l} \sum_{k=1}^{l} a_i x_k 2^{-(i+k)}. \end{aligned} \tag{69}$$

This can be rewritten as a two's-complement number

$$\alpha \cdot x = -c_0 + \sum_{i=1}^{2l} c_i 2^{-i}, \tag{70}$$

where the summation on the right now goes to $2l$, making the length of the product $(2l + 1)$. A two's-complement multiplier usually has one of two means of reducing this product to a $(l + 1)$-bit number. In truncation, all the bits corresponding to $i \geq l + 1$ are discarded. From (70), it is seen that such an operation always decreases the magnitude of a positive number, and increases the magnitude of a negative number, thus introducing a negative bias. In rounding a number, $2^{-(l+1)}$ is added to the product in (70) and the result is truncated to $(l + 1)$ bits. This always selects the $(l + 1)$-bit number that is closest in magnitude to the true product. An ambiguity arises, however, when this product is equidistant from two $(l + 1)$-bit numbers. It is seen from (69) that rounding, in this case, always increases the magnitude of a positive number and decreases the magnitude of a negative number, thus introducing a positive bias in the arithmetic.

It is assumed that the numbers $\pm\alpha$ and x in (69) represent the symbols and the scaled error, respectively, in the updating algorithm. The bias situation in rounding arises when the following conditions occur in (70):

$$c_{l+1} = 1 \quad \text{and} \quad c_i = 0 \quad \text{for} \quad i > l + 1. \tag{71}$$

Inspection of (69) shows that these conditions are equivalent to

$$\sum_{i=1}^{l} \sum_{k=l+1-i}^{l} a_i x_k 2^{-(i+k)} = 2^{-(l+1)}. \tag{72}$$

* It is readily verified that, with this definition, the largest negative number that can be represented is -1, and the largest positive number is $+1 - \text{LSB}$.

The rounding bias can be eliminated by choosing the a_i's in such a manner as to never satisfy (72), regardless of the choice of the x_k's. The search for all of the possible numbers, α, having this property is tedious and will not be pursued here. Rather, two sets of symbols are proposed and shown to have the desired property. The proof requires that the scaled error be reasonably small. More specifically, the magnitude of x must be strictly less than one-half. This assumption is always satisfied in practice, even during start-up.

The first two binary symbols to be considered are defined by

$$\alpha = 0100 \cdots 001 = 2^{-1} + 2^{-l} \tag{73}$$

and

$$-\alpha = 1011 \cdots 111 = -1 + \sum_{i=2}^{l} 2^{-i}. \tag{74}$$

These numbers are equal to $\{\pm((1/2) + \text{LSB})\}$. Considering the positive symbol and replacing the a_i's by their value in (72) yields

$$x_l 2^{-(l+1)} + \sum_{k=1}^{l} x_k 2^{-(l+k)} = 2^{-(l+1)}. \tag{75}$$

This equation can only be satisfied under the condition $x_1 = 1$ and $x_i = 0$ for $i > 1$. From (68), it is seen that the only two numbers satisfying this condition are $\pm 1/2$. These numbers were discarded earlier as valid solutions. For a negative symbol (72) becomes

$$\sum_{i=2}^{l} \sum_{k=l+1-i}^{l} x_k 2^{-(i+k)} = 2^{-(l+1)}. \tag{76}$$

A sequence of x_k's cannot be synthesized to satisfy this equation. The lowest-order bit on the left corresponds to the power $2l$ of 1/2. There is only one term on the left in (76) contributing to this bit's value: $i = k = l$. Since there is no such term on the right, this bit must be zero ($x_l = 0$). The value of second-lowest bit corresponding to a power $(2l-1)$ depends upon two terms: $i = l$, $k = l - 1$ and $i = l - 1$. Therefore, its value is $(x_l + x_{l-1})$, which must be zero modulo-2. Since $x_l = 0$, this implies $x_{l-1} = 0$. Reasoning by induction shows that all of the x_i's must be zero for $i \geq 2$. The highest-order bit corresponds to the $(l + 1)$th power of 1/2. Combining this with the preceding result, we see that only one term contributes to its value; $i = l$, $k = 1$. Therefore, to satisfy (76), x_1 must equal one. As previously noted, the only two numbers that produce the bias situation are $\pm 1/2$, which are not valid solutions.

The second set of symbols that eliminates the multiplier's bias is given by

$$\alpha = 0111 \cdots 111 = 1 - 2^{-l} \tag{77}$$

and

$$-\alpha = 1000 \cdots 001 = -1 + 2^{l}. \tag{78}$$

These numbers are equal to {±(1-LSB)}. Proof by contradiction, similar to the one utilized previously, can be used to show that these symbols eliminate the multiplier's bias. The proof is left as an exercise for the reader.

AUTHORS

John M. Cioffi, B.S. (Electrical Engineering), 1978, University of Illinois; M.S. and Ph.D. (Electrical Engineering), Stanford University, in 1979 and 1984, respectively; Bell Laboratories, 1979–81; IBM Research, 1984—. In 1978 Mr. Cioffi received the Jordan and Davidson awards from the University of Illinois as the outstanding senior in Engineering and the Valedictorian of Engineering, respectively. He is the author of numerous publications in the areas of adaptive filtering and data transmission and storage. Member, Tau Beta Pi, Eta Kappa Nu, Phi Kappa Phi, Phi Eta Sigma, Sigma Xi, IEEE.

Jean-Jacques Werner, Ing. Deg., 1965, INSA, Lyon, France; M.S. (Electrical Engineering), 1967, Laval University, Canada; Eng. Sc.D. (Electrical Engineering), 1973, Columbia University; Bell Laboratories, 1973–1982; AT&T Information Systems, 1983—. At Bell Laboratories and AT&T Information Systems, Mr. Werner has worked on problems in data transmission and digital signal processing. Member, Sigma Xi, IEEE.