A Family of Low-Complexity Blind Equalizers

Cheng-I Hwang and David W. Lin, *Senior Member, IEEE*

***Abstract*—Two important topics in equalizer design are its complexity and its training. We present a family of blind equalizers which, by incorporating a decomposition finite-impulse response filtering technique, can reduce the complexity of the convolution operation therein by about one half. The prototype algorithm in this equalizer family employs the prevalent Godard cost function. Several simplified algorithms are proposed, including a sign algorithm which eliminates multiplications in coefficient adaptation and a few delayed versions. We also study the convergence properties of the algorithms. For the prototype algorithm, we show that, in the limit of an infinitely long equalizer and under mild conditions on signal constellations and channel characteristics, there are only two sets of local minima on the performance surface. One of the sets is undesirable and is characterized by a null equalized channel response. The other corresponds to perfect equalization, which can be reached with proper equalizer initialization. For the simplified algorithms, corresponding cost functions may not exist. Some understanding of their convergence behaviors are obtained via examination of their adaptation equations. Simulation results are presented to demonstrate the performance of the algorithms.**

***Index Terms*—Adaptive filtering, blind equalization, decomposition convolution.**

I. Introduction

In DIGITAL transmission, the problem of intersymbol interference (ISI) is most often mitigated by receiving-end equalization. Two important issues in equalizer design and implementation are its complexity and its training. Regarding complexity, the equalizer is frequently the most complicated element in the receiver. Hence, simplification of the equalizer can reduce the receiver cost effectively. Regarding training, conventional equalizers rely on the use of a training sequence for initial convergence. However, in some situations, it may be costly to send a training sequence, or the training sequence may be unavailable at the receiver. Then blind equalization (or more exactly, blind adaptation) is needed, where the equalizer coefficients are adapted, employing some known statistics of the transmitted data, but not the (unknown) data values. Blind or nonblind, the equalizer switches to decision-directed operation after initial convergence.

Paper approved by R. A. Kennedy, the Editor for Data Communications Modulation and Signal Design of the IEEE Communications Society. Manuscript received August 23, 1999; revised January 21, 2000. This work was supported in part by the National Science Council of the Republic of China under Grant NSC 87-2218-E-009-052.

C.-I Hwang was with the Department of Electronics Engineering and Center for Telecommuications Research, National Chiao Tung University, Hsinchu 30010, Taiwan, R.O.C. He is now with Realtek, Hsinchu 300, Taiwan, R.O.C. (e-mail: alexhuang@realtek.com.tw).

D. W. Lin is with the Department of Electronics Engineering and Center for Telecommunications Research, National Chiao Tung University, Hsinchu 30010, Taiwan, R.O.C. (e-mail: dwlin@cc.nctu.edu.tw).

Digital Object Identifier 10.1109/TCOMM.2004.823588

An adaptive filter contains two major sections: a convolution section and a coefficient adaptation section. Studies on complexity reduction for blind and nonblind algorithms alike have thus far largely concentrated on the coefficient adaptation section. A common technique is to use an adaptation step size which has a simple binary representation. In the case of nonblind adaptive filtering, various "sign-least mean square (LMS) algorithms" have also been studied [1], [3]. Compared to the fundamental LMS algorithms, these sign versions reduce (or totally eliminate) the number of multiplications in coefficient adapation. Taking a different route, Chen *et al.* recently introduced a technique for convolution computation which could reduce the amount of multiplications therein by about 50% [2], [3]. We shall refer to this technique as the decomposition method. Elsewhere, it is sometimes referred to as the algebraic reduction method.

In the case of blind equalization, Godard's method for quadrature amplitude modulation (QAM) signals [4] is a frequently referenced adaptation technique. Due to its carrier independence, the convergence of the equalizer and the carrier recovery circuit can be more easily assured. For complexity reduction, sign algorithms for blind equalization have also been developed [5], [6]. All these blind techniques concern linear equalization. But decision-feedback equalization can be introduced when the equalizer has finished the blind adaptation and switched to decision-directed operation [7].

In this paper, we present several low-complexity blind equalization algorithms. In particular, we incorporate the decomposition convolution technique into blind equalization, and we investigate the equalizer performance both analytically and through computer simulation. The equalizer adaptation technique is based on the Godard $p = 2$ cost function. (A separate paper considers joint operation of the blind equalizer with carrier and timing recovery [7].) Several simplified versions of the resulting algorithm are also derived. Moreover, typical algorithm studies assume that the coefficient adaptation employs the convolution result of the current time. But practical hardware designs are sometimes obliged to use the result of an earlier time, giving rise to the so-called delayed algorithms [8]. We also consider such algorithms.

A main contribution of this paper consists in the discovery of a family of blind equalization algorithms with significantly reduced complexity over their conventional counterparts. The reduction is up to about one half the complexity in the equalizer's convolution operation. Depending on the choice of the equalizer's coefficient adaptation scheme, this is tantamount to between approximately 25% and 50% of the overall equalizer complexity.

The rest of this paper is organized as follows. In Section II, we derive the family of low-complexity blind equalizers. Section III

studies their convergence property. Section IV presents some computer simulation results. And Section V draws the conclusion.

II. Decomposition Blind Equalization Algorithms

We treat the convolution section and the coefficient adaptation section in separate subsections.

A. Decomposition Technique for Convolution

Conventional transversal filters perform convolution as

$$z(n) = \sum_{k=0}^{N-1} x(n-k)w_k(n) \tag{1}$$

where n is the time index, $x(n-k)$ is the filter input, $w_k(n)$ is the kth filter coefficient, and $z(n)$ is the filter output. Chen and Tsay [9] observed that (1) can be rearranged as follows:

$$\begin{aligned} z(n) = &\sum_{k=0}^{\frac{N}{2}-1} [x(n-2k) + w_{2k+1}(n)] \\ &\times [x(n-2k-1) + w_{2k}(n)] \\ &- \sum_{k=0}^{\frac{N}{2}-1} x(n-2k)x(n-2k-1) \\ &- \sum_{k=0}^{\frac{N}{2}-1} w_{2k}(n)w_{2k+1}(n) \end{aligned} \tag{2}$$

where, without significant loss of generality, we have assumed N to be even for convenience. This is the decomposition approach to computing convolution.

In the right-hand side (RHS) of (2), the first term requires $N/2$ multiplications and $(3N/2)-1$ additions at each time. Let $g(n)$ denote the second RHS term in (2). It can be computed recursively as

$$g(n) = g(n-2) + x(n)x(n-1) - x(n-N)x(n-N-1). \tag{3}$$

The computation takes two multiplications and two additions. Alternatively, we can save the value of $x(n)x(n-1)$ in a delay line for later use. Then a multiplication is saved at the cost of some memory space and operations. The last RHS term, in the case of constant-coefficient filters, can be computed at the design stage. In summary, for constant-coefficient filters, the decomposition formulation (2) can obtain the output at each time with on the order of $N/2$ multiplications and $(3N/2)$ additions. Compared with the order of N multiplications and N additions for the conventional method, the decomposition method reduces the amount of multiplications by about 50% at the cost of about a 50% increase in number of additions. But the complexity of digital multiplication can be much greater than that of addition, especially for complex signals such as those encountered in baseband processing of QAM signals. Thus, the decomposition technique offers significant complexity reduction for convolution computation.

In adaptive filtering, however, the coefficients $w_k(n)$ are time varying, and the last RHS term in (2) cannot be computed in advance. A straightforward computation of this term would require $N/2$ multiplications and $N/2-1$ additions at each time, making the computation according to (2) more complicated than the conventional method. To avoid this problem, Chen *et al.* defined a new variable $h(n)$ in its place; i.e., they modified (2) into

$$\begin{aligned} z(n) = &\sum_{k=0}^{\frac{N}{2}-1} [x(n-2k) + w_{2k+1}(n)] \\ &\times [x(n-2k-1) + w_{2k}(n)] \\ &- \sum_{k=0}^{\frac{N}{2}-1} x(n-2k)x(n-2k-1) - h(n) \end{aligned} \tag{4}$$

where the variable $h(n)$ is adapted to estimate $\sum_k w_{2k}(n)w_{2k+1}(n)$. This way, the complexity advantage of the decomposition technique is maintained. Experience shows that, with proper adaptation, $h(n)$ can track $\sum_k w_{2k}(n)w_{2k+1}(n)$ closely. It is this decomposition filtering method that we use in our blind equalizers.

Another complexity issue regarding the decomposition technique is that the power levels of $x(n)$ and $w_k(n)$ may be quite different. Hence, the arithmetic operations in (4) may need to be carried out using greater wordlengths than in the conventional method, offsetting the benefits of reduced multiplications. This problem can be alleviated by scaling the filter input to make the power level of $x(n)$ approximately equal to that of $w_k(n)$, so as to minimize the impact on multiplier size.

Note that, for equalizers, the above equations apply equally well to synchronous and fractionally spaced structures—only let the index n refer to the sample number rather than the symbol number.

B. Coefficient Adaptation

By far, the most common coefficient adaptation algorithms for either blind or nonblind adaptive filtering are of the stochastic gradient type. For this, one typically first defines a cost function (on equalizer output $z(n)$) whose global minimum corresponds to the desired filter coefficient setting. For example, in the case of nonblind adaptive filtering, the most common cost function is the mean-square error (MSE), i.e., $E|z(n) - d_n|^2$, where d_n is the desired filter output. Taking the stochastic gradient of this cost function yields the well-known least mean square (LMS) algorithm. For blind equalization, it is customary to construct cost functions such that the global minima correspond to perfect equalization in the limit of infinitely long equalizers. Consider Godard's cost function [4]. Let $\Psi(z(n))$ denote the cost function and $\Phi(z(n))$ its sample value at time n, that is, $\Psi(z(n)) = E\{\Phi(z(n))\}$. Godard let

$$\Phi(z(n)) = (|z(n)|^p - R_p)^2, \quad R_p = \frac{E|a_n|^{2p}}{E|a_n|^p} \tag{5}$$

where p is some positive integer. Taking the stochastic gradient approach, one obtains the following adaptation equation for equalizer coefficients in the conventional transversal structure:

$$\begin{aligned} w_k(n+1) &= w_k(n) - \mu \nabla_{w_k(n)} \Phi(z(n)) \\ &= w_k(n) - \mu \phi(z(n)) x^*(n-k) \end{aligned} \tag{6}$$

where

$$\begin{aligned}\phi(z(n)) &= \nabla_{z(n)}\Phi(z(n)) \\ &= 2pz(n)|z(n)|^{p-2}(|z(n)|^p - R_p) \end{aligned} \tag{7}$$

and the gradient operator ∇ is defined as $\nabla_c q = (\partial q)/(\partial \Re\{c\}) + j(\partial q)/(\partial \Im\{c\})$. A popular choice for p is $p = 2$. In fact, if $p > 2$, then the dynamic range of $\Phi(z(n))$ becomes too big, degrading the performance.

When the equalizer assumes the decomposition filter structure (4), the stochastic gradient approach yields the following coefficient adaptation formulas:

$$\begin{aligned} w_{2k}(n+1) &= w_k(n) - \mu\nabla_{w_{2k}(n)}\Phi(z(n)) \\ &= w_{2k}(n) - \mu\phi(z(n)) \\ &\quad \times (x(n-2k) + w_{2k+1}(n))^* \end{aligned} \tag{8}$$

$$\begin{aligned} w_{2k+1}(n+1) &= w_k(n) - \mu\nabla_{w_{2k+1}(n)}\Phi(z(n)) \\ &= w_{2k+1}(n) - \mu\phi(z(n)) \\ &\quad \times (x(n-2k-1) + w_{2k}(n))^* \end{aligned} \tag{9}$$

$$\begin{aligned} h(n+1) &= h(n) - \nu\nabla_{h(n)}\Phi(z(n)) \\ &= h(n) + \nu\phi(z(n)) \end{aligned} \tag{10}$$

where μ and ν are the adaptation step sizes. These formulas constitute *the prototype algorithm* in our study.

Now consider the situation where the adaptive equalizer has converged and the equalization is nearly perfect. In this situation, $h(n) \approx \sum_k w_{2k}(n)w_{2k+1}(n)$ and the output $z(n)$ of the decomposition equalizer and that of a conventional transversal equalizer will be about equal. Therefore, the values of $\Phi(z(n))$ from both equalizer structures will be close. By the smoothness of $\Phi(z(n))$, we expect that the values of $\nabla_{w_k(n)}\Phi(z(n))$ for both structures be also similar. Hence, (8) and (9) should behave similar to (6). And we may consider using (6) in place of (8) and (9), which results in a simplified algorithm for coefficient adaptation as

$$w_k(n+1) = w_k(n) - \mu\phi(z(n))x^*(n-k) \tag{11}$$

$$h(n+1) = h(n) + \nu\phi(z(n)). \tag{12}$$

We term this the *reduced algorithm*. Simulation results indicate that, interestingly, the convergence behavior of the reduced algorithm is statistically similar to that of the original, even if used from the beginning, i.e., not just after convergence.

For further complexity reduction, we consider the following *sign algorithm* to reduce the number of multiplications:

$$w_k(n+1) = w_k(n) - \mu\,\mathrm{sgn}(\phi(z(n)))x^*(n-k) \tag{13}$$

$$h(n+1) = h(n) + \nu\phi(z(n)) \tag{14}$$

where the signum function for a complex number c is defined as $\mathrm{sgn}(c) = \mathrm{sgn}(\Re\{c\}) + j\mathrm{sgn}(\Im\{c\})$. The adaptation equation for $h(n)$ is not changed, since it does not contain multiplications to start with (except for a scaling by ν), and since taking the sign of $\phi(z(n))$ would make the adaptation of $h(n)$ too coarse. We remark that the sign algorithm in [6] employs a different $\phi(z(n))$, given by $\phi(z(n)) = \mathrm{sgn}(z(n))[|\Re\{z(n)\}| + |\Im\{z(n)\}| - R]$ for some R, which is derived from a modified cost function of that of Godard, and hence, is not exactly the sign version of Godard's algorithm.

The above algorithms assume that the coefficient adaptation employs the convolution result of the current time. As mentioned previously, hardware-efficient designs are sometimes obliged to use the result of an earlier time, giving rise to delayed algorithms such as the *delayed reduced algorithm*

$$w_k(n+1) = w_k(n) - \mu\phi(z(n-D))x^*(n-(k+D)) \tag{15}$$

$$h(n+1) = h(n) + \nu\phi(z(n-D)). \tag{16}$$

Replacing the $\phi(z(n-D))$ in (15) by $\mathrm{sgn}(\phi(z(n-D)))$ yields the *delayed sign algorithm.*

III. Convergence Properties

Convergence behaviors of blind equalizers have so far defied a complete analysis. Researchers often resort to an examination of the performance surface associated with infinitely long equalizers under the given cost function to get some grasp. For conventional transversal equalizers, Godard studied the stationary points of his $p = 2$ cost function, and showed that the global minima correspond to perfect equalization [4]. Foschini further demonstrated that all other stationary points are either local maximum or saddle points, and hence, the global minima are the only attractors [10]. Convergence behaviors of finite-length blind equalizers are more complicated, and there may exist local minima [11]. For simplicity, we only consider the limiting case of infinitely long equalizers.

We first examine the stationary points associated with the prototype algorithm. Following a somewhat similar route to that of Godard [4], we show that perfect equalization defines local minima, which are outperformed only by another set of local minima whose cost value is exactly zero, and which correspond to a completely null equalized channel response. The derivation also appears in a more condensed form in [7]. Then, by studying the eigenvalue structure of the Hessian matrices of the cost function at the stationary points, we show that all other stationary points are either local maximum or saddle points. Therefore, with proper initialization and suitable choice of the adaptation step sizes, the algorithm can approach perfect equalization.

Next, we consider the reduced and the sign algorithms. Unfortunately, we are unable to show whether there exist corresponding cost functions. Thus, for these algorithms, the notion of performance surface is (as yet) ill defined, and we are unable to conduct a convergence analysis based on study of the performance surface. But some understanding of the convergence behaviors can still be attained by examining the coefficient adaptation equations. And the analysis of the average convergence properties also applies to the delayed algorithms, although they have different dynamic behaviors.

A. Stationary Points of the Prototype Algorithm

Let s_k denote the convolution of the channel impulse response with w_k. Then, in the case of a noiseless channel, the decomposition equalizer output $z(n)$ is given by

$$z(n) = \sum_k s_k a_{n-k} - h + \sum_k w_{2k}w_{2k+1} \tag{17}$$

where a_n denotes the channel input at time n, and we have omitted the time indexes from h and w_k to signify the fact that we are interested in the performance surface defined by these quantities, rather than their time variation. For our study, parametrization of $\Psi(z(n))$ in terms of s_k or in terms of w_k makes no difference if the channel response is invertible, for, in this case, w_k and s_k are related by a nonsingular linear transform. Hence, although the stationary points of the cost function appear in different locations in the two spaces, their curvature types (minimum, maximum, saddle, etc.) are not changed.

Treat $z(n)$ as a function of s_k and h. A stationary point on the performance surface corresponding to the Godard $p = 2$ cost function satisfies $\nabla_h \Psi(z(n)) = 0$ and $\nabla_{s_l} \Psi(z(n)) = 0 \; \forall l$. The former leads to

$$E[(|z(n)|^2 - R_2)z(n)] = 0 \tag{18}$$

whereas the latter

$$E[(|z(n)|^2 - R_2)(z(n)a^*_{n-l} - \Re\{z(n)\}\nabla_{s_l}\Re\{\hat{h}\} - \Im\{z(n)\}\nabla_{s_l}\Im\{\hat{h}\})] = 0 \tag{19}$$

where, for convenience, we have defined

$$\hat{h} = h - \sum_k w_{2k} w_{2k+1}. \tag{20}$$

By (18), (19) can be simplified to

$$E[(|z(n)|^2 - R_2)z(n)a^*_{n-l}] = 0. \tag{21}$$

Consider the situation of perfect equalization, that is, $\hat{h} = 0, s_l \neq 0$ for some l, and $s_k = 0 \; \forall k \neq l$. Substituting these conditions into (18) and (21), we see that a sufficient condition for the equalities to hold is that $E(a) = 0$ and $E(|a|^2 a) = 0$, which is true for common square QAM constellations with equal symbol probabilities. Thus, perfect equalization defines stationary points. We now prove that these points are local minima.

Some manipulation of (21) yields

$$s_l \left\{ [E|a_n|^4 - 2(E|a_n|^2)^2]|s_l|^2 + 2(E|a_n|^2)^2 \sum_k |s_k|^2 - E|a_n|^4 + 2|\hat{h}|^2 E|a_n|^2 \right\} = 0 \tag{22}$$

where, in addition to $E(a) = 0$ and $E(|a|^2 a) = 0$, we have assumed that a_n is independent and identically distributed (i.i.d.), and that $E(a^2) = 0$. (The last condition holds for common square QAM constellations with equal symbol probabilities.) Therefore, at any stationary point, if there are M nonzero elements in the sequence $\{s_k\}$, then all these elements have the same magnitude.

On the other hand, (18) leads to

$$-2\hat{h}E|a_n|^2 \sum_k |s_k|^2 - \hat{h}|\hat{h}|^2 + \hat{h}R_2 = 0 \tag{23}$$

under the same set of conditions on a_n. Therefore, either

$$\hat{h} = 0 \tag{24}$$

or

$$|\hat{h}|^2 = -2E|a_n|^2 \sum_k |s_k|^2 + \frac{E|a_n|^4}{E|a_n|^2}. \tag{25}$$

If $\hat{h} = 0$, then the situation is equivalent to that of a common transversal equalizer already dealt with by Godard, who showed that the cost function value is given by ([4, Eq. (34)])

$$\Psi_{1,M}(z(n)) = \frac{(E|a_n|^4)^2}{(E|a_n|^2)^2} - \frac{M(E|a_n|^4)^2}{E|a_n|^4 + 2(M-1)(E|a_n|^2)^2}. \tag{26}$$

It is easily seen that $\Psi_{1,0}(z(n)) > \Psi_{1,M}(z(n)) \; \forall M > 0$. Godard proved that, for $M > 0$, the value increases with M if $E|a_n|^4 < 2(E|a_n|^2)^2$, a condition satisfied by common square QAM constellations. Hence, $\Psi_{1,M}(z(n))$ attains its minimum at $M = 1$, i.e., at perfect equalization. Moreover, at perfect equalization, we have $|s_l| = 1$ ([4, Eq. (32)]). On the other hand, if $\hat{h}$ satisfies (25), then substituting it into (22), we obtain

$$|s_l|^2 = \frac{E|a_n|^4}{2(M+1)(E|a_n|^2)^2 - E|a_n|^4} \quad \forall s_l \neq 0 \tag{27}$$

and thus

$$|\hat{h}|^2 = \frac{E|a_n|^4[2(E|a_n|^2)^2 - E|a_n|^4]}{E|a_n|^2[2(M+1)(E|a_n|^2)^2 - E|a_n|^4]}. \tag{28}$$

The corresponding cost function value is then given by

$$\Psi_{2,M}(z(n)) = \frac{M(E|a_n|^4)^2}{2(M+1)(E|a_n|^2)^2 - E|a_n|^4}. \tag{29}$$

It is straightforward to show that, if $2(E|a_n|^2)^2 > E|a_n|^4$, then $\Psi_{2,M}(z(n)) > \Psi_{1,M}(z(n))$ for all $M > 0$. Unfortunately, $\Psi_{2,0}(z(n)) = 0$, which obviously gives the absolute minimum of the cost function. Hence, the global minima of the performance surface correspond to a completely null equalized channel impulse response. The desired condition of perfect equalization corresponds to stationary points whose cost value is the second lowest among all stationary points.

Fortunately, as shown immediately below, the second-lowest stationary points that correspond to perfect equalization are local minima. Further, it is shown in the Appendix that, in fact, all other stationary points except that corresponding to null filtering are either local maximum or saddle points. Thus, convergence to perfect equalization can be attained by a stochastic gradient algorithm with suitable initialization and proper choice of the adaptation step sizes.

To show that perfect equalization corresponds to local minima, note from the Godard cost function (5) with $p = 2$ that we have

$$\begin{aligned}\Psi(z(n)) = {} & 2(E|a_n|^2)^2\left(\sum_k |s_k|^2\right)^2 \\ & -[2(E|a_n|^2)^2 - E|a_n|^4]\sum_k |s_k|^4 \\ & -2E|a_n|^4\sum_k |s_k|^2 \\ & +|\hat{h}|^4 + 4E|a_n|^2\sum_k |s_k|^2|\hat{h}|^2 \\ & -2R_2|\hat{h}|^2 + R_2^2. \end{aligned} \tag{30}$$

It is then not difficult to show that, at perfect equalization, any perturbation in h and $|s_k|$ would lead to an increase in the value of $\Psi(z(n))$. Thus, perfect equalization defines local minima.

Although the above derivation has assumed zero-mean data, slight modification of the algorithm will make it capable of operating under nonzero-mean data. Transmission of nonzero-mean data may arise, for example, when it is desirable to send a pilot tone at the carrier frequency. Such a tone may be effected by adding a direct current (DC) bias to an otherwise symmetric baseband data constellation. In conventional transversal equalizers, such a bias in data may pose a problem, because it may result in a particularly large eigenvalue in the equalizer's input autocorrelation matrix, to the detriment of its adaptation [12]. A way to handle this problem in conventional adaptive filters is to add a DC tap to estimate and cancel the DC bias. In the case of our prototype blind equalizer algorithm, let the mean of the transmitted data be m_a, and define $\hat{a}_n = a_n - m_a$. Then $\hat{a}_n$ will have zero mean, and we have

$$z(n) = \sum_k s_k\hat{a}_{n-k} - h + \sum_k w_{2k}w_{2k+1} + \sum_k s_k m_a. \tag{31}$$

To address the situation of a nonzero m_a, all we need to do is to redefine

$$R_2 = \frac{E|\hat{a}_n|^4}{E|\hat{a}_n|^2}. \tag{32}$$

It can be shown that the stationary points of $\Psi(z(n))$ again exhibit the above-discussed properties, with the value of h at the second-lowest minima now equal to $\sum_k w_{2k}w_{2k+1} + \sum_k s_k m_a$ to take care of the effect of nonzero m_a automatically. Therefore, the decomposition equalizer easily handles data with nonzero DC.

B. Simplified Algorithms

As noted before Section III-A, cost functions corresponding to the simplified algorithms are not found and may not exist. We attempt to gain some understanding of the convergence behaviors by examining the coefficient adaptation equations.

Consider first the reduced algorithm. It is not difficult to show that the average behavior of the algorithm is similar to the prototype algorithm when h is converged, thus providing additional justification to the reduction, in addition to that discussed in the last section. For this, we see from (10) that when h is fully converged in the prototype algorithm, we should have $E\{\phi(z(n))\} = 0$. From (7) with $p = 2$, this is exactly the stationarity condition (18). If the adaptation step size μ in (8) and (9) is small, then we may presume that $w_k(n)$ is only weakly dependent on $z(n)$, and thus

$$\begin{aligned} & E\{\phi(z(n))(x(n-2k) + w_{2k+1}(n))^*\} \\ & \quad \approx E\{\phi(z(n))x^*(n-2k)\} \\ & \qquad + E\{\phi(z(n))\}E\{w_{2k+1}^*(n)\} \\ & \quad = E\{\phi(z(n))x^*(n-2k)\} \end{aligned} \tag{33}$$

$$\begin{aligned} & E\{\phi(z(n))(x(n-2k-1) + w_{2k}(n))^*\} \\ & \quad \approx E\{\phi(z(n))x^*(n-2k-1)\} \\ & \qquad + E\{\phi(z(n))\}E\{w_{2k}^*(n)\} \\ & \quad = E\{\phi(z(n))x^*(n-2k-1)\}. \end{aligned} \tag{34}$$

Therefore, the average behavior of the reduced algorithm is similar to the prototype algorithm when h is converged.

For the sign algorithm, we noted earlier that our algorithm is different from that in [6], in that the former is a simplification of the prototype algorithm, while the latter has a well-defined cost function different from Godard's. Concerning the convergence behaviors, a peculiar property of our sign algorithm is that, unlike the prototype and the reduced algorithms, perfect equalization does not necessarily define equilibrium points (where the average adaptations of the equalizer coefficients are zero). To see this, let c_k denote the original channel impulse response. Then $x(n) = \sum_k c_k a_{n-k}$. Now consider a situation of perfect equalization, where $z(n) = a_{n-l}$ for some l. The expected value of coefficient adjustment in (13), disregarding the scaling by μ, is given by

$$\begin{aligned} & E\{\mathrm{sgn}(\phi(z(n)))x^*(n-k)\} \\ & \quad = E\{\mathrm{sgn}(\phi(a_{n-l}))x^*(n-k)\} \\ & \quad = E\{\mathrm{sgn}(a_{n-l}(|a_{n-l}|^2 - R_2))x^*(n-k)\} \\ & \quad = E_l\{E_{k\neq l}[\mathrm{sgn}(a_{n-l}(|a_{n-l}|^2 - R_2))x^*(n-k)|a_{n-l}]\} \\ & \quad = E_l\{\mathrm{sgn}(a_{n-l}(|a_{n-l}|^2 - R_2))c_{l-k}^* a_{n-l}^*\} \\ & \quad = c_{l-k}^* \cdot E\{\mathrm{sgn}(|a_{n-l}|^2 - R_2)[\mathrm{sgn}(a_{n-l})a_{n-l}^*]\} \end{aligned} \tag{35}$$

where $E_{k\neq l}$ denotes expectation over all a_{n-k} with $k \neq l$, E_l denotes expectation over a_{n-l}, and we have assumed a_n to be zero-mean and i.i.d. The RHS in (35) can be interpreted as follows.

The quantity $\mathrm{sgn}(|a_{n-l}|^2 - R_2)$ serves to attach different signs to $\mathrm{sgn}(a_{n-l})a_{n-l}^*$, according to whether a_{n-l} lies within a circle of radius $\sqrt{R_2}$ around the origin on the complex plane. The expectation, therefore, classifies the points in the constellation of a_{n-l} into these two categories, takes the sum of $\mathrm{sgn}(a_{n-l})a_{n-l}^*$ for each category, obtains their difference, and divides the result by the constellation size of a_{n-l}. Now, the quantity $\mathrm{sgn}(a_{n-l})a_{n-l}^*$ is a negation of the phase of a_{n-l} through complex conjugation, followed by a rotation in angle by $(\pi/4) + (\pi/2)(q-1)$ (and a scaling of the magnitude by $\sqrt{2}$) through multiplication by $\mathrm{sgn}(a_{n-l})$, where q is the quadrant a_{n-l} lies in. By this, the constellation of a_{n-l} is folded into a segment spanning between $\pm(\pi/4)$ in phase about the real axis on the complex plane, as illustrated in Fig. 1 for

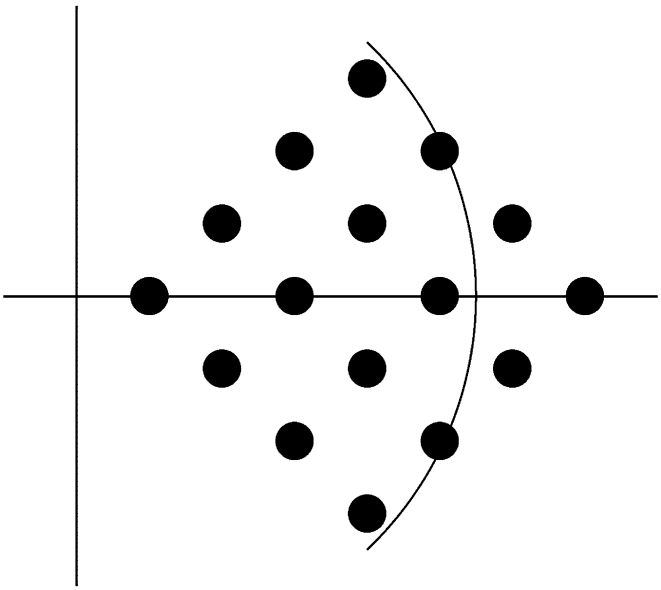


Fig. 1. Folded constellation of 64-QAM given by $\mathrm{sgn}(a_n)a_n^*$.

the common square 64-QAM constellation. In the figure, we have also sketched an arc at $\sqrt{R_2}$ from the origin. Now, if the original constellation of a_{n-l} is vertically symmetric (i.e., symmetric about the real axis), then the folded constellation is also vertically symmetric, resulting in a null imaginary part upon taking the expectation in (35). However, it is straightforward to show that the real parts of $\mathrm{sgn}(|a_{n-l}|^2 - R_2)[\mathrm{sgn}(a_{n-l})a_{n-l}^*]$ do not average to zero for a number of common constellations, such as the 16-, 64-, and 256-QAM, no matter whether we let $\mathrm{sgn}(0)$ be equal to $+1, -1$, or 0. We also see that this condition cannot be rectified by using a different value for R_2, because, for these constellations, no value of R_2 can make the real parts average to zero. Therefore, perfect equalization does not define equilibrium points for the sign algorithm in these cases. Rather, residual ISI or noise must be present in $z(n)$ at an equilibrium point. As blind equalization usually serves only to facilitate the initial convergence of an equalizer (to open the eye) before switching to decision-directed operation, the above peculiarity does not pose a problem, as long as it does not prevent such convergence. This is indeed the case observed in our simulation.

IV. Simulation Results

As mentioned, blind equalization usually serves only to facilitate the initial convergence of an equalizer. We consider a two-stage blind-adaptation procedure as follows: 1) adaptation with a larger step size for fast convergence; and 2) adaptation with a smaller step size to reduce the residual ISI. Then the equalizer enters decision-directed operation. As this simulation study is geared at investigating the performance of the blind equalization algorithms, we limit ourselves to consideration of linear equalization only, even in the stage of decision-directed operation. For simplicity, we let the channel be noise free. This is not expected to degrade the usefulness of the results, for as we shall see, there is already an amount of residual ISI in the

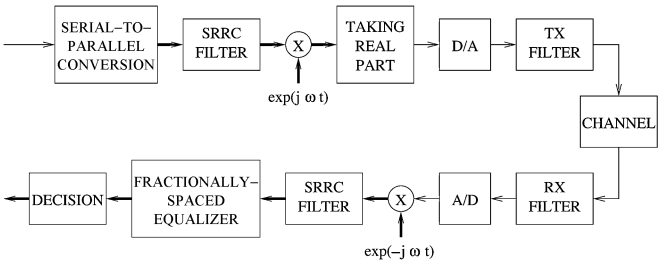


Fig. 2. Block diagram of the transmission system, where thin arrows indicate real signals and thick arrows indicate complex signals.

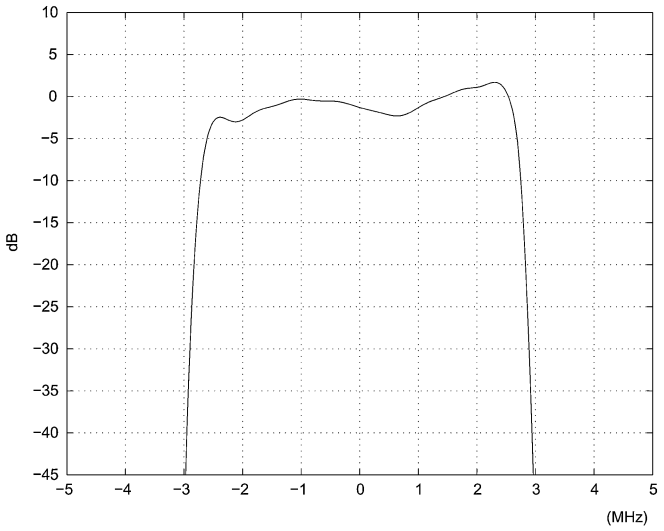


Fig. 3. Magnitude response of the simulated channel.

equalizer output. Such residual ISI already plays the role of additive noise.

We simulate digital coaxial-cable transmission using 64- and 256-QAM signals. The overall transmission system is as shown in Fig. 2. The channel bandwidth is 6 MHz and the signaling rate is 5.38 Mbaud. The two square-root raised-cosine (SRRC) filters each contain 64 taps. The magnitude response of the simulated (equivalent baseband) channel, including the two SRRC filters, is as shown in Fig. 3. The equalizer is $T/2$-spaced with 16 "linear taps" (i.e., the w_k) plus the bias tap (i.e., h). It is initialized to have $w_7 = 1$. The input to the equalizer is manually scaled ("gain controlled") to have a peak-to-peak magnitude variation of approximately 16 (resp. 32) in the two quadrature channels for 64-QAM (resp. 256-QAM), so that in the absence of ISI, the minimum distance between neighboring signal points would be approximately two. The quantity R_2 is also defined based on a minimum distance of two between nearest neighbors in each signal constellation. Each stage of blind adaptation contains 30 000 symbols. For convenience, the adaptation step sizes will be described in vector form, with the ith element giving the step size employed in the ith stage of equalizer operation.

First, consider 64-QAM with zero-mean data. We first examine the performance of the prototype algorithm and the reduced algorithm. Fig. 4 compares the MSE convergence of these algorithms with that of the Godard algorithm, where the adaptation step sizes in the two blind-adaptation stages are set somewhat arbitrarily to $[2^{-22}, 2^{-24}]$ for the linear taps for all the al-

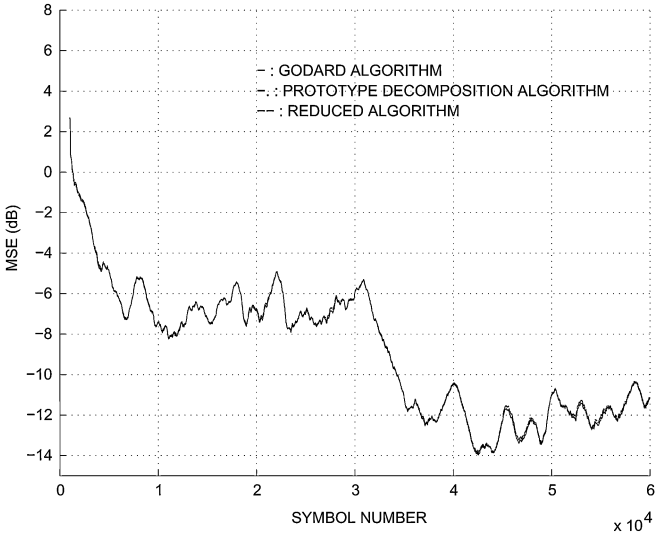


Fig. 4. MSE convergence of the Godard algorithm, the prototype algorithm, and the reduced algorithm under 64-QAM.

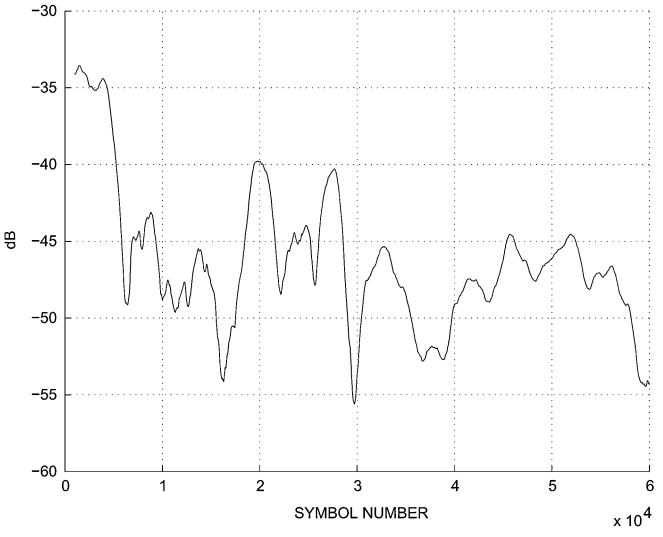


Fig. 5. Convergence of the new variable h under 64-QAM in the reduced algorithm. Curve shows $10\log_{10}\overline{|h-\sum_k w_{2k}w_{2k+1}|^2}/|\sum_k w_{2k}w_{2k+1}|^2$ where the overlines denote time averages.

gorithms, and that for the bias tap set to $[2^{-19}, 2^{-19}]$ for the prototype and the reduced algorithms. The MSE is computed by averaging over 1000 symbols and plotted one point per 10 symbols. We see that the three algorithms have almost identical convergence characteristics. Hence, the reduced algorithm, being the simplest of all, is the preferred choice. Therefore, in the following, we shall be concerned primarily with the reduced algorithm and its simplified versions. Fig. 5 depicts the convergence curve of h for the reduced algorithm, which shows that h converges to $\sum_k w_{2k}w_{2k+1}$, as desired.

We now examine the performance of two simplified versions of the reduced algorithm, namely, the delayed algorithm (with $D = 1$) and the (undelayed) sign algorithm. Fig. 6 plots the MSE convergence of these algorithms together with that of the reduced algorithm. The adaptation step sizes are the same as given previously, except for the linear taps in the sign algorithm, which are $[2^{-14}, 2^{-16}]$. These step sizes have been chosen to make the sign algorithm's convergence speed visually approximately the same as the other two. Smaller step sizes can lower the steady-state MSE of the sign algorithm, but at the price of a slower convergence. From the figure, we see that, in this case, one-symbol delayed adaptation does not degrade the performance, while, as mentioned, it may have some hardware benefit. Note also that, since the most important purpose of blind equalization is to cut the decision-point error to the extent that equalizer convergence in decision-directed operation can be ensured, some degradation in the steady-state performance of the blind-adaptation stage can be accepted if such convergence is not hampered. On this ground, the sign algorithm has a complexity advantage, as it does not need multiplications in coefficient adaptation.

Before considering the combined use of delayed and sign adaptation, let us briefly look into the effects of input–output scaling and that of nonzero DC in data. Recall that the purpose of equalizer input–output scaling is to approximately equalize the peak signal magnitude in the equalizer delay line and the peak equalizer tap values, to attain some hardware complexity merit. We have observed that the maximum magnitude of the equal-

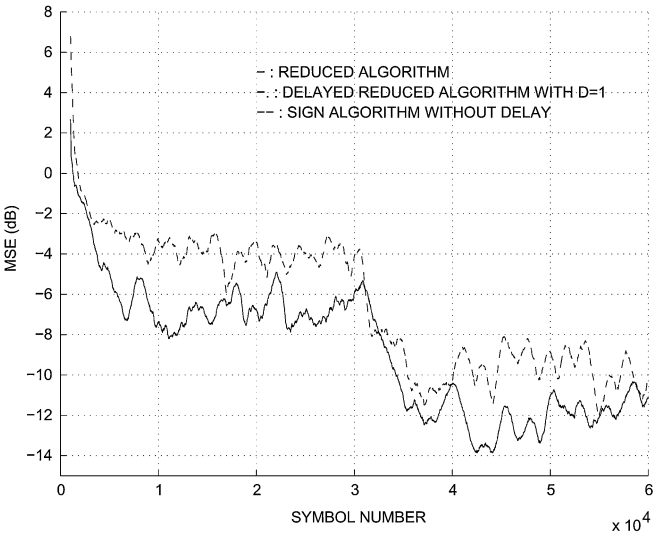


Fig. 6. MSE convergence of several simplified algorithms under 64-QAM.

izer tap values lies around 0.5–1. Therefore, a 16-fold downscaling of the equalizer input, which has been approximately normalized to a maximum magnitude of eight in both quadrature branches as noted previously, will achieve the purpose. Simulation results show that the input–output scaling leads to a fractional-decibel penalty in MSE, compared with the basic reduced algorithm, but the convergence curve is otherwise quite similar. For algorithm performance in nonzero-mean data, we consider a DC bias of 0.5 in both the quadrature branches of the transmitted QAM symbols. The modified target power level of (32) is used. Simulation results show that the MSE convergence is similar to the case of zero-mean data.

We now consider the combined use of delayed and sign adaptation and input–output scaling. Fig. 7 compares the convergence of the blind-adapted delayed sign algorithm (with $D = 1$) with input–output signal scaling (by 16) with that of the training-sequence-based sign LMS algorithm. The latter employs the conventional transversal filter structure. In

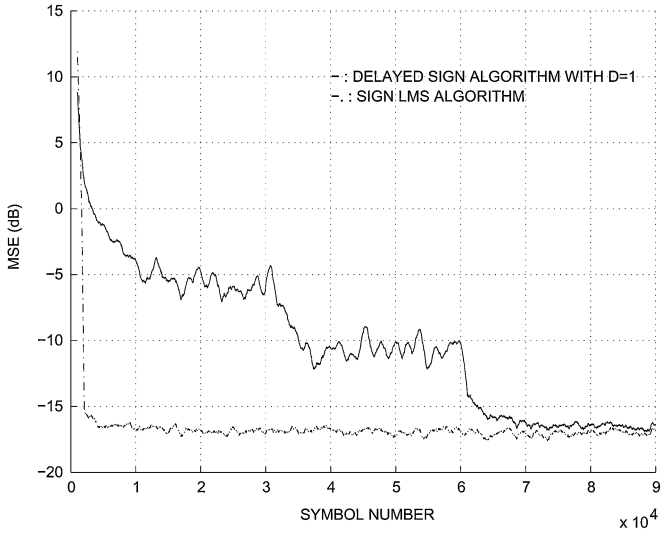


Fig. 7. MSE convergence under 64-QAM of the training-sequence-based conventional sign LMS equalizer under zero-mean data and the blind-adapted delayed sign decomposition equalizer with input–output scaling under nonzero-mean data.

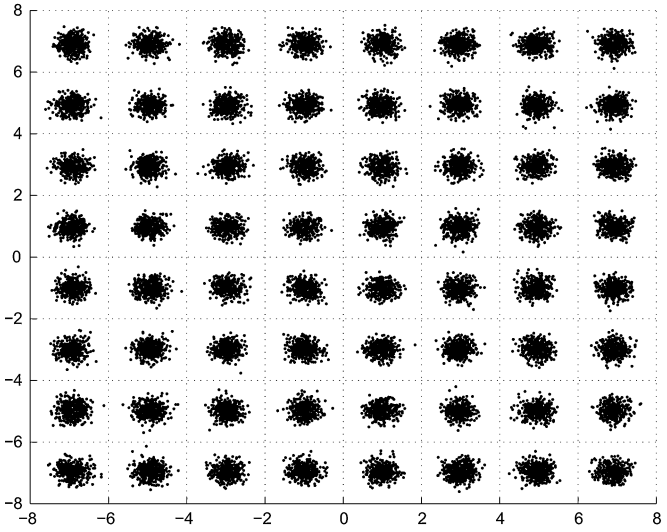


Fig. 8. Equalizer output signal constellation after convergence in stage-2 blind adaptation of the delayed sign algorithm with input–output scaling under 64-QAM with nonzero mean.

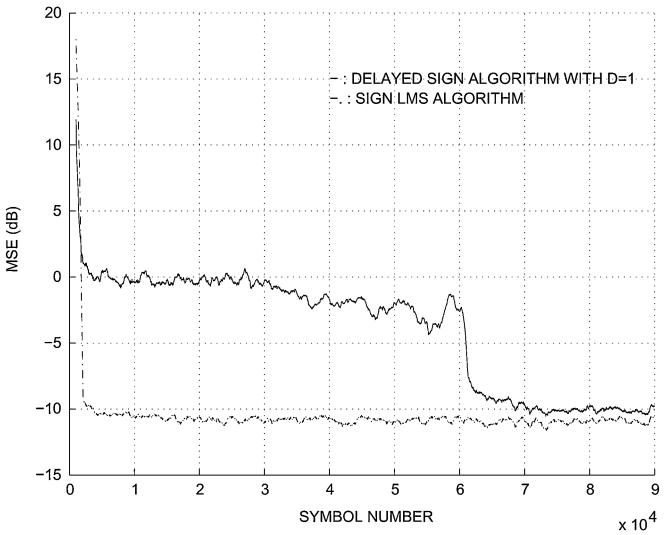


Fig. 9. MSE convergence under 256-QAM of the training-sequence-based conventional-sign LMS equalizer under zero-mean data and the blind-adapted delayed sign decomposition equalizer with input–output scaling under nonzero-mean data.

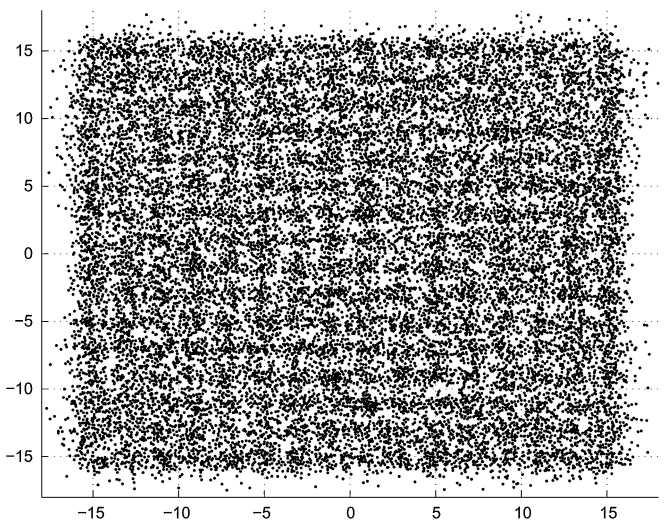


Fig. 10. Equalizer output signal constellation in blind-adaptation stage 2, after 5000 symbols into it, of the delayed sign algorithm with input–output scaling under 256-QAM with nonzero mean.

the blind-adapted case, a DC bias of 0.5 is present in both quadrature branches of the transmitted data, while in the case of training-sequence-based sign LMS, the data have zero mean. Decision-directed operation is entered after 60 000 symbols for blind-adapted equalization and after 30 000 symbols for training-sequence-based equalization. The adaptation step sizes employed in the blind-adapted equalization are $[2^{-11}, 2^{-13}, 2^{-11}]$ for the linear taps and $[2^{-17}, 2^{-19}, 2^{-11}]$ for h (where the third vector elements are the step sizes for decision-directed operation). These adaptation step sizes have been chosen heuristically to achieve relatively fast convergence, while yielding a low steady-state MSE. The step size for the training-sequence-based sign LMS is 2^{-11} throughout. We see from Fig. 7 that, although the MSE values of the delayed sign algorithm during blind adaptation are larger than that of the conventional sign LMS during training-sequence-based adaptation, the final performance in decision-directed operation is comparable. Fig. 8 shows the equalizer output constellation after convergence in stage two of blind adaptation. The eyes are clearly open.

A simulation of 256-QAM transmission is shown in Fig. 9, where the simulation conditions are the same as for Fig. 7, except those noted below. First, the input–output scaling in the delayed sign algorithm is by a factor of 32, due to the different input scale. And second, the adaptation step sizes for the blind-adapted equalization are $[2^{-10}, 2^{-13}, 2^{-11}]$ for the linear taps and $[2^{-18}, 2^{-20}, 2^{-11}]$ for h. These step sizes have been chosen, likewise, to achieve relatively fast convergence while yielding a relatively low steady-state MSE. The step size for the training-sequence-based sign LMS is again 2^{-11} throughout. Observe that the blind-adapted equalizer again converges under such conditions and ends in similar MSE performance as the

training-sequence-based conventional sign LMS in decision-directed operation. Fig. 10 shows the equalizer output constellation in stage two of blind adaptation after 5000 symbols into it. The eyes are not as cleanly open as in the case of 64-QAM, but they can be discerned.

V. Conclusion

We presented a family of blind equalization algorithms which employ a decomposition FIR filtering technique to carry out the convolution therein. By this, the complexity of convolution can be reduced to about one half. A stochastic-gradient-based prototype algorithm for equalizer coefficient adaptation was derived. The algorithm can handle both zero-mean and nonzero-mean data by varying an algorithm parameter which controls the expected output magnitude of the equalizer. Several simplified algorithms were also obtained, including a reduced algorithm which behaves similarly to the prototype algorithm at near-perfect equalization, a sign algorithm which eliminates multiplication in equalizer coefficient adaptation, and several delayed versions thereof.

We studied the convergence properties of the algorithms. For the prototype algorithm, we showed that in the limit of an infinitely long equalizer and under mild conditions on signal constellations and channel characteristics, there are only two sets of local minima on the performance surface. One of the sets is undesirable and corresponds to completely null equalized channel responses. The other set corresponds to perfect equalization. Other stationary points are either local maximum or saddle points. Thus, convergence to perfect equalization can be obtained with suitable initialization. For the simplified algorithms, some understanding of their convergence behaviors were obtained via examination of their adaptation equations.

We presented some simulation results to demonstrate the performance of the algorithms. In these examples, setting one of the equalizer's central tap to unity made it converge successfully.

For future work, a subject of interest is the theoretical analysis of the convergence properties of the prototype algorithm assuming finite-length equalizers. For the simplified algorithms, it would be interesting to show definitely if corresponding cost functions (or approximate ones) exist. It would also be interesting to consider the performance of the proposed algorithms on fading multipath channels.

Appendix
Curvature at the Stationary Points

The eigenvalue structure of the Hessian matrix at a stationary point bears information on the type of stationarity there. For notational convenience, let $\overline{q}$ denote the expected value of a quantity q, and let c_r and c_i denote the real and the imaginary parts, respectively, of a complex quantity c. Though somewhat tedious, it is not difficult to derive the following second-order derivatives (where $\alpha, \beta = r$ or i):

$$\frac{\partial^2 \Psi(z(n))}{\partial h_\alpha^2} = 4\left(2\sum_k |s_k|^2\overline{|a|^2} + |\hat{h}|^2 - R_2\right) + 8\hat{h}_\alpha^2 \tag{36}$$

$$\frac{\partial^2 \Psi(z(n))}{\partial h_r \partial h_i} = \frac{\partial^2 \Psi(z(n))}{\partial h_i \partial h_r} = 8\hat{h}_r\hat{h}_i \tag{37}$$

$$\begin{aligned}\frac{\partial^2 \Psi(z(n))}{\partial s_{l\alpha}^2} &= 4\left[|s_l|^2\overline{|a|^4} + 2\sum_{k\neq l}|s_k|^2\overline{|a|^2}^2 + (2|\hat{h}|^2 - R_2)\overline{|a|^2}\right] + 8s_{l\alpha}^2\overline{|a|^4} \\ &\quad + 16 s_{l\alpha}\overline{|a|^2}\frac{\partial|\hat{h}|^2}{\partial s_{l\alpha}} + 2[2\overline{|a|^2}\sum_k |s_k|^2 \\ &\quad + |\hat{h}|^2 - R_2]\frac{\partial^2|\hat{h}|^2}{\partial s_{l\alpha}^2} + 2\left[\frac{\partial|\hat{h}|^2}{\partial s_{l\alpha}}\right]^2\end{aligned} \tag{38}$$

$$\begin{aligned}\frac{\partial^2 \Psi(z(n))}{\partial s_{lr}\partial s_{li}} &= \frac{\partial^2 \Psi(z(n))}{\partial s_{li}\partial s_{lr}} \\ &= 8 s_{lr}s_{li}\overline{|a|^4} + 8\overline{|a|^2}\left[s_{lr}\frac{\partial|\hat{h}|^2}{\partial s_{li}} + s_{li}\frac{\partial|\hat{h}|^2}{\partial s_{lr}}\right] \\ &\quad + 2\frac{\partial|\hat{h}|^2}{\partial s_{lr}}\frac{\partial|\hat{h}|^2}{\partial s_{li}} + 2\left[2\overline{|a|^2}\sum_k |s_k|^2 + |\hat{h}|^2 - R_2\right] \\ &\quad \times \frac{\partial^2|\hat{h}|^2}{\partial s_{lr}\partial s_{li}}\end{aligned} \tag{39}$$

$$\begin{aligned}\frac{\partial^2 \Psi(z(n))}{\partial s_{l\alpha}\partial h_\beta} &= -4\frac{\partial\hat{h}_\beta}{\partial s_{l\alpha}}\left[2\sum_k |s_k|^2\overline{|a|^2} + |\hat{h}|^2 - R_2\right] \\ &\quad -4\hat{h}_\beta\left[4 s_{l\alpha}\overline{|a|^2} + \frac{\partial|\hat{h}|^2}{\partial s_{l\alpha}}\right]\end{aligned} \tag{40}$$

$$\begin{aligned}\frac{\partial^2 \Psi(z(n))}{\partial s_{l\alpha}\partial s_{m\beta}} &= 16 s_{l\alpha}s_{m\beta}\overline{|a|^2}^2 \\ &\quad +8\overline{|a|^2}\left[s_{l\alpha}\frac{\partial|\hat{h}|^2}{\partial s_{m\beta}} + s_{m\beta}\frac{\partial|\hat{h}|^2}{\partial s_{l\alpha}}\right] \\ &\quad + 2\frac{\partial|\hat{h}|^2}{\partial s_{l\alpha}}\frac{\partial|\hat{h}|^2}{\partial s_{m\beta}} + 2\left[2\overline{|a|^2}\sum_k |s_k|^2 \right. \\ &\quad \left. + |\hat{h}|^2 - R_2\right]\frac{\partial^2|\hat{h}|^2}{\partial s_{l\alpha}\partial s_{m\beta}}.\end{aligned} \tag{41}$$

Before proceeding, we note that most of the Hessian matrices at the stationary points will be found to possess zero eigenvalues. However, since the Godard cost (30) contains no odd powers of $\hat{h}$ or $s_k\ \forall k$, no stationary point can be reflexive along the direction of any Hessian eigenvector associated with a zero eigenvalue. Thus, a stationary point can only be a local extremum or a saddle point.

A. Case Where $\hat{h} = 0$

When $M > 1$, let s_l and s_m be two nonzero equalized channel coefficients. Then from (36)–(41), the Hessian submatrix associated with $\hat{h}_r, \hat{h}_i, s_{lr}, s_{li}, s_{mr}$, and s_{mi} is given by

$$\begin{aligned} &16\overline{|a|^2}^2[\underline{0}_2', \underline{s}_l', \underline{s}_m']' \text{ [same]} \\ &\quad -8(2\overline{|a|^2}^2 - \overline{|a|^4})[\underline{0}_2', \underline{s}_l', \underline{0}_2']' \text{ [same]} \\ &\quad -8(2\overline{|a|^2}^2 - \overline{|a|^4})[\underline{0}_2', \underline{0}_2', \underline{s}_m']' \text{ [same]} \\ &\quad +4\left(2\overline{|a|^2}\sum_k |s_k|^2 - R_2\right) \\ &\quad \times [-I_2, (\nabla_l \hat{h})', (\nabla_m \hat{h})']' \text{ [same]} \end{aligned} \tag{42}$$

where $\underline{0}_2$ is the two-element zero vector, I_2 is the 2×2 identity matrix, $\underline{s}_k = [s_{kr}, s_{ki}]' \; \forall k$, and

$$\nabla_k \hat{h} = \begin{bmatrix} \frac{\partial \hat{h}_r}{\partial s_{kr}} & \frac{\partial \hat{h}_i}{\partial s_{kr}} \\ \frac{\partial \hat{h}_r}{\partial s_{ki}} & \frac{\partial \hat{h}_i}{\partial s_{ki}} \end{bmatrix} \tag{43}$$

$\forall k$. Note that all the vectors in the four outer products in (42) are pairwise linearly independent. Note also that the coefficient of the second outer product is negative when $2\overline{|a|^2}^2 > \overline{|a|^4}$. Thus, the Hessian submatrix has at least one negative eigenvalue. Therefore, the overall Hessian matrix at the stationary point is not nonnegative definite ([13, Sec. X.4]), and the point cannot be a local minimum. In fact, the Hessian matrix possesses both positive and negative eigenvalues. Consequently, the stationary point is a saddle point.

When $M = 1$, let s_l be the nonzero equalized channel coefficient and consider any N zero equalized channel coefficients $s_{m_n}, n = 1, \ldots, N$. Then the Hessian submatrix associated with $\hat{h}$, s_l, and s_{m_n}, $n = 1, \ldots, N$, is given by

$$\begin{aligned} &8\overline{|a|^4}\,\mathrm{diag}\,(O_2, \underline{s}_l\underline{s}_l', O_2, O_2, \ldots, O_2) \\ &\quad +4(2\overline{|a|^2}^2 - \overline{|a|^4})\,\mathrm{diag}(O_2, O_2, I_2, I_2, \ldots, I_2) \\ &\quad +\frac{4(2\overline{|a|^2}^2 - \overline{|a|^4})}{\overline{|a|^2}} \\ &\quad \times [-I_2, (\nabla_l \hat{h})', (\nabla_{m_1} \hat{h})', (\nabla_{m_2} \hat{h})', \ldots, (\nabla_{m_N} \hat{h})']' \\ &\quad \cdot \text{ [same]} \end{aligned} \tag{44}$$

where O_2 is the 2×2 zero matrix and we have used the fact that $|s_l| = 1$ ([4, Eq. (32)]). We see that the coefficients of all three matrix terms are positive, provided $2\overline{|a|^2}^2 > \overline{|a|^4}$. Thus, the Hessian submatrix is nonnegative definite, and since N is arbitrary, so is the overall Hessian matrix at this stationary point. Therefore, the stationary point is a local minimum.

When $M = 0$, again consider any N zero-equalized channel coefficients $s_{m_n}, n = 1, \ldots, N$. The Hessian submatrix associated with $\hat{h}$ and these coefficients is given by

$$\begin{aligned} &-4\overline{|a|^4}\,\mathrm{diag}\,(O_2, I_2, I_2, \ldots, I_2) \\ &\quad -4R_2[-I_2, (\nabla_{m_1} \hat{h})', (\nabla_{m_2} \hat{h})', \ldots, (\nabla_{m_N} \hat{h})']' \text{ [same]}. \end{aligned} \tag{45}$$

As the coefficients of the two matrix terms are both negative, the Hessian submatrix is nonpositive definite and so is the overall Hessian matrix at this point. Hence, the point is a local maximum.

B. Case Where $\hat{h} \neq 0$

The condition $M = 0$ corresponds to the undesired global minima. For $M > 0$, let s_l be a nonzero equalized channel coefficient. The Hessian submatrix associated with s_l is given by

$$2\left(4\overline{|a|^2}\underline{s}_l + \begin{bmatrix} \frac{\partial |\hat{h}|^2}{\partial s_{lr}} \\ \frac{\partial |\hat{h}|^2}{\partial s_{li}} \end{bmatrix}\right) (\text{[same]})' - 8(4\overline{|a|^2}^2 - \overline{|a|^4})\underline{s}_l\underline{s}_l' \tag{46}$$

from which we can calculate its determinant as

$$-16(4\overline{|a|^2}^2 - \overline{|a|^4})\left(s_{lr}\frac{\partial |\hat{h}|^2}{\partial s_{li}} - s_{li}\frac{\partial |\hat{h}|^2}{\partial s_{lr}}\right)^2. \tag{47}$$

Thus, the Hessian submatrix is not nonnegative definite. Hence, the overall Hessian matrix is not nonnegative definite ([13, Sec. X.4]), and this stationary point cannot be a local minimum. By considering Hessian submatrices which include derivatives with respect to $\hat{h}_r$ and $\hat{h}_i$, and using an eigenvalue argument as that done previously for the condition $M > 1$ under the case $\hat{h} = 0$, we can show that the overall Hessian matrix possesses both positive and negative eigenvalues, and hence, the stationary point is a saddle point.

REFERENCES

[1] M. G. Bellanger, *Adaptive Digital Filters and Signal Analysis*. New York: Marcel Dekker, 1987.

[2] S.-G. Chen, Y.-A. Kao, and K.-Y. Tsai, "A new efficient LMS adaptive filtering algorithm," *IEEE Trans. Circuits Syst. II*, vol. 43, pp. 372–378, May 1996.

[3] C.-I Hwang, T.-C. Tang, D. W. Lin, and S.-G. Chen, "An efficient FSE/DFE-based HDSL equalizer with new adaptive algorithms," in *Proc. IEEE Int. Conf. Communications*, vol. 1, 1994, pp. 288–292.

[4] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication system," *IEEE Trans. Commun.*, vol. COM-28, pp. 1867–1875, Nov. 1980.

[5] V. Weerackody, S. A. Kassam, and K. R. Laker, "Sign algorithms for blind equalization and their convergence analysis," *Circuits Syst. Signal Processing*, vol. 10, no. 4, pp. 394–431, 1991.

[6] V. Weerackody, S. A. Kassam, and K. R. Laker, "A simple hard-limited adaptive algorithm for blind equalization," *IEEE Trans. Circuits Syst. II*, vol. 39, pp. 482–485, July 1992.

[7] C.-I Hwang and D. W. Lin, "Joint low-complexity blind equalization, carrier recovery, and timing recovery with application to cable modem transmission," *IEICE Trans. Commun.*, vol. E82-B, pp. 120–128, Jan. 1999.

[8] G. Long, F. Ling, and J. G. Proakis, "The LMS algorithm with delayed coefficient adaptation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 1397–1405, Sept. 1989.

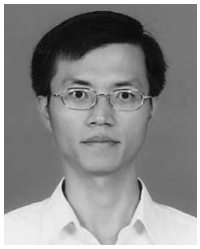
[9] S.-G. Chen and R. J. Tsay, "A novel fast-filtering algorithm and its implementation," in *Proc. EUSIPCO*, Brussels, Belgium, Aug. 1992, pp. 945–948.

[10] G. J. Foschini, "Equalizing without altering or detecting data," *AT&T Tech. J.*, vol. 64, no. 8, pp. 1885–1911, Oct. 1985.

[11] Z. Ding, R. A. Kennedy, B. D. O. Anderson, and C. R. Johnson, Jr., "Ill-convergence of Godard blind equalizers in data communication systems," *IEEE Trans. Commun.*, vol. 39, pp. 1313–1327, Sept. 1991.

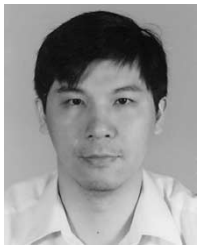
[12] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[13] F. R. Gantmacher, *The Theory of Matrices*. New York: Chelsea, 1977.



Cheng-I Hwang received the B.S. and Ph.D. degrees in electronics engineering from National Chiao Tung University, Hsinchu, Taiwan, R.O.C., in 1991 and 1999, respectively.

He was with the Industrial Technology Research Institute, Hsinchu, Taiwan from 1999 to 2003. He is currently a Project Manager with Realteck Semiconductor Corporation, Hsinchu, Taiwan, R.O.C. His main research area is digital communication systems, particularly cable modems and digital TV.



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

He was with Bell Laboratories during 1981–1983, and with Bellcore during 1984–1990 and again during 1993–1994. Since 1990, he has been a Professor in the Department of Electronics Engineering and the Center for Telecommunications Research, National Chiao Tung University, except for a leave during 1993–1994. He has conducted research in digital adaptive filtering and telephone echo cancellation, digital subscriber line and coaxial network transmission, speech and video coding, and wireless communication. His research interests include various topics in communication engineering and signal processing.