Blind Fractionally Spaced Equalization of Noisy FIR Channels: Direct and Adaptive Solutions

Georgios B. Giannakis, *Fellow, IEEE*, and Steven D. Halford, *Student Member, IEEE*

***Abstract*— Blind fractionally spaced equalizers reduce intersymbol interference using second-order statistics without the need for training sequences. Methods for finding FIR zero-forcing blind equalizers directly from the observations are described, and adaptive versions are developed. In contrast, most current methods require channel estimation as a first step to estimating the equalizer. The direct methods can be zero-forcing, minimum mean-square error, or even minimum mean square error (MMSE) within the class of zero-forcing equalizers. Performance of the proposed methods and comparisons with existing approaches are shown for a variety of channels, including an empirically measured digital microwave channel.**

***Index Terms*— Adaptive equalization, cyclostationarity, fractional sampling.**

I. Introduction and Problem Statement

IN HIGH-SPEED digital communications, the channel often introduces memory in the received signal, which in turn spreads the symbols over time. This spreading induces a distortion known as intersymbol interference (ISI), which, in order to maintain reliable performance, must be removed at the receiver by equalization. Since, in practical systems, the channel is not known *a priori*, training sequences can be used to estimate the channel and find the necessary equalizer. Alternatively, *blind* equalizers (see, e.g., [3]) exploit knowledge about the structure of the input (e.g., whiteness) in conjunction with the outputs in order to estimate the equalizer. Since they do not require extra bandwidth for training, blind equalizers have received great research and practical interest, and many methods have been proposed [25], [27].

Traditionally, higher (than second-) order statistics of the symbol rate sampled outputs are used to either estimate the channel first and then calculate the equalizer (c.f., [25]) or to directly estimate the equalizer (c.f., [27]). More recently, Tong *et al.* [31] showed that second-order statistics of the outputs contain sufficient information to estimate most communication channels when the outputs are sampled faster than the symbol rate (fractionally sampled). Based on the seminal work in [31], many effective blind methods have been proposed for estimating the channel from the output-only second-order statistics [1], [19], [22], [32], [35]. Each of these methods provides a blind estimate of the channel that can then be used to find the maximum-likelihood (ML) estimate of the transmitted sequence using the Viterbi algorithm [26, p. 588]. For a given channel estimate, the ML approach provides the minimum probability of error estimate, although it can be computationally intensive. Liu and Xu [21] have proposed a least-squares method that, while suboptimal to ML, directly estimates the sequence from the observed data. Similarly, Tong [30] has recently proposed a method that, although not ML, uses the Viterbi algorithm to directly estimate the source sequence without explicitly estimating the channel. A third approach, which is similar in philosophy to the higher order methods of [27], is to bypass the channel estimation step and directly estimate a linear filter that can remove the ISI and/or suppress the additive noise [11], [28]. While also suboptimal to the ML approach, the direct estimation of the equalizer is computationally efficient and lends itself easily to the development of adaptive methods for tracking time-varying channels.

In this paper, we present three methods for finding linear equalizers directly from the data. Each method is derived from a specific performance criterion:

a) zero-forcing (ZF) or ISI removal;
b) minimum mean-square error (MMSE);
c) minimum mean-square error within the class of ZF.

The novelties of this work over the existing direct equalizer of [28] and the subsequent work in [29] are

i) explicit inclusion of additive noise;
ii) one-step computation of the blind MMSE equalizer;
iii) direct computation of the ZF for any delay;
iv) hybrid method to find the "optimal" ZF equalizer for a given delay and length.

In addition, we develop adaptive methods that allow for the tracking of time-varying channels. In comparison with the adaptive approach of [8], the methods presented here are computationally simpler at the expense of some design flexibility. Finally, the methods of this paper have been extended to include deterministically modeled inputs [9], nonlinear channels [12], two-dimensional (2-D) image models [10], and multiuser communications [33].

Manuscript received May 1, 1995; revised February 19, 1997. This work supported by ONR Grant N00014-93-0485. Part of the work in this paper was presented at the International Conference on Acoustics, Speech, and Signal Processing, Detroit, MI, May 9–12, 1995 and at the Conference on Information Sciences and Systems, Baltimore, MD, March 22–24, 1995. The associate editor coordinating the review of this paper and apporving it for publication was Dr. Jitendra Tugnait.

G. B. Giannakis is with the Department of Electrical Engineering, University of Virginia, Charlottesville, VA 22903-2442 USA.

S. D. Halford is now with the Government Communications Systems Division, Harris Corporation, Melbourne, FL USA.

Publisher Item Identifier S 1053-587X(97)05784-X.

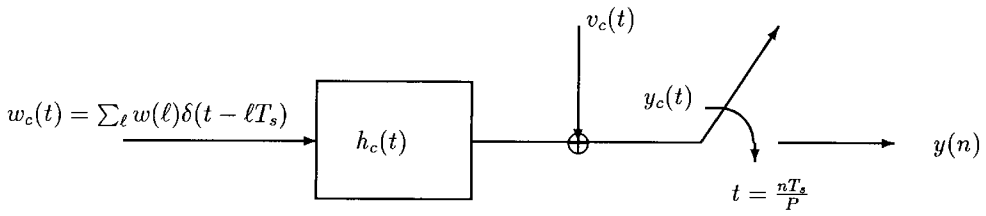

Fig. 1. Fractionally sampled communication system.

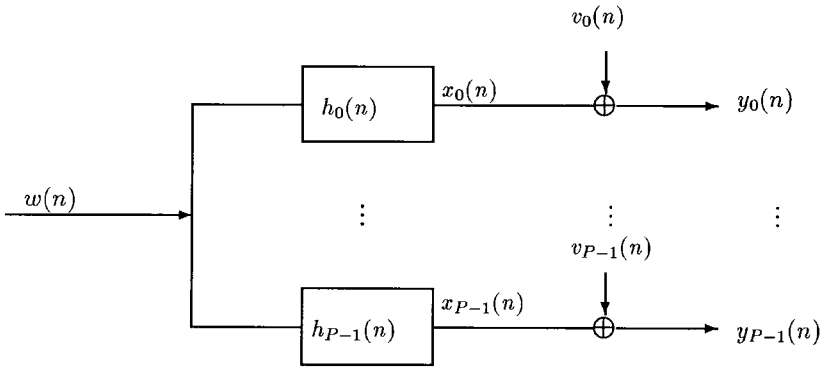

Fig. 2. Multichannel model for FS channel.

II. Mathematical Framework

Consider the continuous-time *fractionally sampled* communication system shown in Fig. 1. The sequence

- $w(n)$ denotes information symbols;
- T_s symbol duration;
- $h_c(t)$ the "composite" channel;
- $v_c(t)$ additive noise thta is assumed to be stationary as well as uncorrelated with $w(n)$;
- P an integer denoting the amount of oversampling.

The constituents of the composite channel $h_c(t)$ include the known transmit and receive filters as well as the unknown transmission channel. For this system, the signal at the sampler is $y_c(t) = \int_{-\infty}^{\infty} w_c(\tau)h_c(t-\tau)\,d\tau + v_c(t) = \Sigma_\ell\, w(\ell)h_c(t - \ell T_s) + v_c(t)$. If $y_c(t)$ is sampled at $t = nT_s/P$, the received data are $y(n) := y_c(t)$ for $t = nT_s/P$, which is $y(n) = \Sigma_\ell\, w(\ell)h_c(nT_s/P - \ell T_s) + v_c(nT_s/P)$. If we consider the discrete-time sequences $w(n)$ and $y(n)$ as the input and output, respectively, it is convenient to write the input–output relationship as an equivalent discrete-time system

$$y(n) = \sum_{\ell=-\infty}^{\infty} w(\ell)h(n - \ell P) + v(n) = x(n) + v(n) \quad (1)$$

where $h(n)$ and $v(n)$ are the discrete-time equivalents of $h_c(t)$ and $v_c(t)$, respectively, and $x(n)$ is the discrete-time equivalent of the noise-free received signal. As will be seen, the outputs described by (1) have a periodically time-varying correlation (with period P). In many cases, periodically correlated signals are conveniently represented by vector stationary processes [24]. Upon defining $y_i(n) := y(nP + i)$, the single-input, single-output (SISO) relationship of (1) accepts an *equivalent* single-input, multiple-output (SIMO) description as given by (see also Fig. 2)

$$y_i(n) = \sum_{\ell=-\infty}^{\infty} w(\ell)h_i(n - \ell) + v_i(n) \quad \text{for } i = 0, 1, \cdots, P-1 \quad (2)$$

where $h_i(n) := h(nP + i)$, and $v_i(n) := v(nP + i)$. We note that a length P block of the time-varying, scalar output of (1) is now represented as a time-invariant, $P \times 1$ vector output. More specifically, by concatenating equations from (2) for $i \in [0, P-1]$, we can represent $y(n)$ in a vector form as

$$\boldsymbol{y}(m) = \sum_{\ell} w(\ell)\boldsymbol{h}(m - \ell) + \boldsymbol{v}(m) \quad (3)$$

where

$$\begin{aligned}\boldsymbol{y}(m) &:= [y(mP) \quad y(mP-1) \cdots y(mP-P+1)]' \\ &= [y_o(m) y_1(m) \cdots y_{P-1}(m)]', \boldsymbol{h}(m) \\ &:= [h_0(m) \quad h_1(m) \cdots h_{P-1}(m)]', \boldsymbol{v}(m) \\ &:= [v_0(m) \quad v_1(m) \quad \cdots \quad v_{P-1}(m)]'\end{aligned}$$

and $'$ denotes transpose.

If $h_c(t)$ is FIR of order $L_h T_s$ (i.e., $h_c(t) \neq 0$ for $t \in [0, L_h T_s]$), it follows that the subchannels $\{h_i(n)\}_{i=0}^{P-1}$ will be of order L_h. Correspondingly, for N vector observations

$$\boldsymbol{y}_N(n) := [\boldsymbol{y}'(n) \quad \boldsymbol{y}'(n-1) \quad \cdots \quad \boldsymbol{y}'(n-N+1)]' \quad (4)$$

only input symbols $w(m)$ for $m = n, n-1, \cdots, n-N-L_h+1$ will affect the observed data [see (3)]. Collecting equations from (3) with $m = n, n-1, \cdots, n-N+1$, we rewrite the

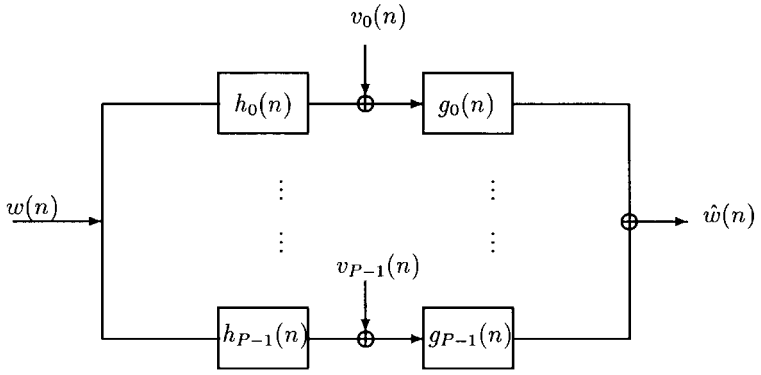

Fig. 3. Linear equalization for FS channels.

input–output input–output relation as (see also [22])

$$\boldsymbol{y}_N(n) = \mathcal{H}\boldsymbol{w}_{N+L_h}(n) + \boldsymbol{v}_N(n) \tag{5}$$

where $\mathcal{H}$ is the $NP \times (N + L_h)$ block Toeplitz matrix

$$\mathcal{H} := \begin{bmatrix} \boldsymbol{h}(0) & \boldsymbol{h}(1) & \cdots & \boldsymbol{h}(L_h) & \cdots & 0 \\ 0 & \boldsymbol{h}(0) & \cdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & \boldsymbol{h}(L_h-1) & \boldsymbol{h}(L_h) \end{bmatrix} \tag{6}$$

and

$$\boldsymbol{w}_{N+L_h}(n) := [w(n) \quad w(n-1) \quad \cdots \quad w(n-N-L_h+1)]' \tag{7}$$

$$\boldsymbol{v}_N(n) := [\boldsymbol{v}'(n) \quad \boldsymbol{v}'(n-1) \quad \cdots \quad \boldsymbol{v}'(n-N+1)]' \tag{8}$$

are $(N + L_h) \times 1$ and $NP \times 1$ vectors, respectively. In the following section, we describe the covariance of $\boldsymbol{y}_N(n)$, which will be useful for finding the equalizers directly from the output.

Remark 1: In general, L_h is not known *a priori*. As in [22], [28], [31], and [35], we will assume in the sequel that at least an estimate of L_h has been obtained. Statistical methods for estimating L_h are given in [15], and rank-based methods can be found in [35].

A. Cyclic Statistics

Consider the correlation $c_{2y}(n;m) := E\{y(n) \quad y^*(n+m)\}$ of the scalar output in (1)

$$c_{2y}(n;m) = \sum_{\ell_1}\sum_{\ell_2} c_{2w}(\ell_2-\ell_1)h(n-\ell_1 P)h^*(n-\ell_2 P) + c_{2v}(m) \tag{9}$$

where $c_{2w}(m) := E\{w(n)w^*(n+m)\}$ and $c_{2v}(m) := E\{v(n) \quad v^*(n+m)\}$ are the correlations of the stationary input and noise, respectively, and where $*$ represents complex conjugation. From (9), it is straightforward to verify that the correlation is periodically time-varying in n with period P (i.e., $c_{2y}(n;m) = c_{2y}(n+\ell P;m) \forall$ integer ℓ). Similar to (9), the correlation of the vector $\boldsymbol{y}_N(n)$ in (4) is given by

$$\begin{aligned} \boldsymbol{C}_{2y}^{(N)} &:= E\{\boldsymbol{y}_N(n) \quad \boldsymbol{y}_N(n)^{*'}\} \\ &= \mathcal{H}\boldsymbol{C}_{2w}^{(N+L_h)}\mathcal{H}^{*'} + \boldsymbol{C}_{2v}^{(N)} \end{aligned} \tag{10}$$

where $\boldsymbol{C}_{2w}^{(N+L_h)} := E\{\boldsymbol{w}_{N+L_h}(n) \quad \boldsymbol{w}_{N+L_h}^{*'}(n)\}$, $\boldsymbol{C}_{2v}^{(N)} := E\{\boldsymbol{v}_N(n) \quad \boldsymbol{v}_N^{*'}(n)\}$, and $^{*'}$ indicates conjugate transpose (Hermitian).

B. Zero-Forcing Equalizers

Consider the FIR linear equalizer shown in Fig. 3, where $\{g_i(n)\}_{i=0}^{P-1}$ for $i = 0, 1, \cdots P-1$ is the order L_g equalizer of the ith subchannel. Ideally, the impulse response coefficients (or taps) of $g_i(n)$ should be chosen to minimize the probability of error between decisions based on the equalizer output $\hat{w}(n)$ and the true values $w(n)$. Unfortunately, the probability of error is a nonlinear function of the equalizer taps, and alternative criteria are necessary. In the absence of noise, one natural choice is to require $\hat{w}(n-d) = w(n)$ for some integer delay d. This type of equalizer is known as zero-forcing (ZF) [26, ch. 10]. More precisely, a ZF equalizer whose subchannels are order L_g is described by

$$\sum_{i=0}^{P-1}\sum_{\ell=0}^{L_h} h_i(\ell)g_i^{(d)}(m-\ell) = \delta(m-d) \tag{11}$$

where superscript (d) refers to the delay $d \in [0, L_h + L_g]$. Choosing $N = L_g + 1$ in (6) and defining $\boldsymbol{H} := \mathcal{H}|_{N=L_g+1}$ implies that $\boldsymbol{H}'$ is

$$\boldsymbol{H}' := \begin{bmatrix} \boldsymbol{h}'(0) & \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \boldsymbol{h}'(1) & \boldsymbol{h}'(0) & \mathbf{0} & \cdots & \mathbf{0} \\ \boldsymbol{h}'(2) & \boldsymbol{h}'(1) & \boldsymbol{h}'(0) & \ddots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{h}'(L_h) & \boldsymbol{h}'(L_h-1) & \cdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \cdots & \cdots & \boldsymbol{h}'(L_h-1) \\ \mathbf{0} & \cdots & \cdots & \cdots & \boldsymbol{h}'(L_h) \end{bmatrix}.$$

This definition allows the ZF condition to be written in matrix form as

$$\boldsymbol{H}'\boldsymbol{g}_d = \boldsymbol{e}_{d+1} \tag{12}$$

where $\boldsymbol{g}_d := [\boldsymbol{g}_d'(0) \quad \boldsymbol{g}_d'(1) \quad \cdots \quad \boldsymbol{g}_d'(L_g)]'$ is an $(L_g+1)P \times 1$ vector of the equalizer taps corresponding to delay $d, \boldsymbol{g}_d'(\ell) := [g_0^{(d)}(\ell) \quad g_1^{(d)}(\ell) \quad \cdots \quad g_{P-1}^{(d)}(\ell)]$, and

$$\boldsymbol{e}_{d+1} := [\underbrace{0 \quad 0 \quad \cdots \quad 0}_{d \text{ zeros}} \quad 1 \quad 0 \quad \cdots \quad 0]'$$

is a $(L_h+L_g+1)\times 1$ vector with a 1 as the $(d+1)$st element and zeros elsewhere. We address the existence of $\boldsymbol{g}_d$ satisfying (12) in the following theorem (see also [28] for $P=2$).

Theorem 1—Existence and Uniqueness of ZFE: Assuming the subchannels $\{h_i(n) i = 0,1,\cdots,P-1\}_{n=0}^{L_h}$ have no common roots, an FIR ZF equalizer with subchannels of order L_g exists, provided $L_g \geq L_h - 1$. The ZF equalizer is unique when $L_g = L_h/(P-1) - 1$, provided $L_g \geq L_h - 1$, or alternatively, when the minimum norm solution is adopted for solving (12).

Proof: The vector $\boldsymbol{e}_{d+1}$ belongs to a complex space of dimension $L_g + L_h + 1$. In order for a solution $\boldsymbol{g}_d$ of (12) to exist, the range space for the columns of $\boldsymbol{H}'$ must be of dimension $L_g + L_h + 1$. Since $\boldsymbol{H}'$ has exactly $L_g + L_h + 1$ rows, the range space of $\boldsymbol{H}'$ will be of dimension $L_g + L_h + 1$ if and only if (iff) $\boldsymbol{H}'$ is full-row rank. As in [5], [22], [28], and [31], the block Toeplitz structure of $\boldsymbol{H}'$ implies that it will be full row rank iff there are no common zeros among the subchannels, and $L_g \geq (L_h) - 1$. When $L_g = (L_h/(P-1)) - 1$ and $L_g \geq (L_h) - 1$, then $\boldsymbol{H}'$ is a full-rank, square matrix, and $\boldsymbol{g}_d$ can be found uniquely from (12). □

We note that the requirement of no common zeros is shared with all second-order-based blind methods. Tugnait [34] discusses a class of channels that never satisfy this condition. The condition on L_g implies that ZF equalizers must be of a minimum length. We note that this result was also shown by Slock [28] for the $P=2$ case in the context of ZF FS equalizers. It is a direct consequence of the Bezout identity [20, p. 382] (see also [13]).

In the case of $P(L_g+1) > L_g + L_h + 1$ (i.e., the matrix is "fat"), the ZF equalizer is not unique, and additional constraints are needed to solve (12). One possible constraint to (12) is the minimum norm, which leads to the pseudo-inverse solution (e.g., [18, p. 410]), i.e.,

$$\boldsymbol{g}_d = \boldsymbol{H}^*(\boldsymbol{H}'\boldsymbol{H}^*)^\dagger \boldsymbol{e}_{d+1} = (\boldsymbol{H}')^\dagger \boldsymbol{e}_{d+1} \tag{13}$$

where † indicates pseudoinverse. In the sequel, we will exploit the fact that ZF equalizers are not necessarily unique in order to find equalizers that are ZF and have desirable noise suppression characteristics (Section III-C).

III. Direct Blind Equalizers

In this section, we describe three methods for finding equalizers directly from the set of observations $\boldsymbol{y}_N(n)$ in (4). Section III-A presents a method for finding a ZF equalizer [see (12)], whereas Section III-B presents a method for obtaining the minimum mean-square error blind equalizer. In Section III-C, we describe a method that finds the ZF equalizer with the minimum MSE. We conclude in Section III-D with a discussion of estimation from sample values.

A. ZF Equalizers

Since ZF equalizers are optimized to suppress ISI without regard to noise, we first consider the noise-free correlation matrix of (10) for $N = L_g + 1$

$$\boldsymbol{C}_{2x}^{(L_g+1)} := \boldsymbol{C}_{2y}^{(L_g+1)} - \boldsymbol{C}_{2v}^{(L_g+1)} = \boldsymbol{H}\boldsymbol{C}_{2w}^{(L_h+L_g+1)}\boldsymbol{H}^{*'}. \tag{14}$$

In the sequel, we will drop the "size" notation from the correlation matrices unless specifically required, e.g., $\boldsymbol{C}_{2y} := \boldsymbol{C}_{2y}^{(L_g+1)}$. As in [28], [29], and [32], we assume the input $w(n)$ is i.i.d., and therefore, $\boldsymbol{C}_{2w} = \sigma_w^2 \boldsymbol{I}$, where $\boldsymbol{I}$ is the identity matrix. In this case, the noise-free output correlation matrix is

$$\boldsymbol{C}_{2x} = \sigma_w^2 \boldsymbol{H}\boldsymbol{H}^{*'}. \tag{15}$$

Taking the complex conjugate and multiplying on the right by the zero-delay $(d=0)$ ZF equalizer yields [see (12)]

$$\boldsymbol{C}_{2x}^* \boldsymbol{g}_0 = \sigma_w^2 \boldsymbol{H}^* \boldsymbol{e}_1 = \sigma_w^2 \boldsymbol{H}^*(:,1) \tag{16}$$

where $\boldsymbol{H}(:,1)$ denotes the first column (MATLAB notation) of matrix $\boldsymbol{H}$. This results leads to the following lemma.

Lemma 1—Direct Equalizer: Under the conditions of Theorem 1, the zero-delay $(d=0)$ equalizer can be found to within a scale ambiguity of $\sigma_w^2 h^*(0)$ directly from the correlation matrix by solving

$$\begin{aligned} \boldsymbol{C}_{2x}^* \boldsymbol{g}_0 &= \sigma_w^2 [\boldsymbol{h}^{*'}(0) \quad 0 \quad \cdots \quad 0]' \\ &= \sigma_w^2 [h^*(0) \quad 0 \quad \cdots \quad 0]'. \end{aligned} \tag{17}$$

Proof: Consider first using the pseudo-inverse to find $\boldsymbol{g}_0$ from (17). We note that $\boldsymbol{C}_{2x}^{*\dagger} = \boldsymbol{H}'\dagger \boldsymbol{H}^{*\dagger}$ so that $\boldsymbol{g}_0 = \sigma_w^2 \boldsymbol{H}'^{\dagger} \boldsymbol{H}^{*\dagger} \boldsymbol{H}^*(:,1)$. Since $\boldsymbol{H}^*(:,1) = \boldsymbol{H}^* \boldsymbol{e}_1$

$$\boldsymbol{g}_0 = \sigma_w^2 \boldsymbol{H}'^{\dagger} \boldsymbol{H}^{*\dagger} \boldsymbol{H}^* \boldsymbol{e}_1 = \sigma_w^2 \boldsymbol{H}'^{\dagger} \boldsymbol{e}_1.$$

Compared with (13), the $d=0$ ZF equalizer can be found directly from the correlation matrix of the data by solving (17). When $\boldsymbol{C}_{2x}$ is square $(L_g = (L_h)/(P-1) - 1)$, the pseudo-inverse is replaced by a matrix inverse, and the ZF equalizer found by solving (17) is unique. The ambiguity arises when $\sigma_w^2 h^*(0)$ is not known *a priori*, and we assume $\sigma_w^2 h^*(0) = 1$. □

While the ZF equalizer removes all the ISI (ideally), it does not suppress the noise; rather, the noise is now colored and possibly enhanced by the filter $\{g_i^{(d)}(n): i = 0,1,\cdots P-1\}_{n=0}^{L_g}$ (c.f., [26, p. 606]). The degree of noise enhancement and even the variance of the equalizer estimates may depend on the delay d of the ZF equalizer. In many situations, therefore, a ZF equalizer with nonzero delay is desirable (for specific examples, see [6]). Given the zero-delay ZF equalizer, it is possible to use (11) to generate a system of equations that can then be used to solve for $\{h_i(n): i = 0,1,\cdots,P-1\}_{n=0}^{L_h}$. Once $h(n)$ is obtained, (12) can be used to find $\boldsymbol{g}_d$. This approach, however, requires two additional matrix inverses

after solving for $\boldsymbol{g}_0$. Surprisingly, it is possible to find the delay d ZF equalizer from the zero delay $(d=0)$ equalizer without an additional matrix inverse, as we show next.

First, we make the following definitions for notational convenience:

$$\tilde{\boldsymbol{C}}_{2x} := \boldsymbol{C}_{2x}^{(2L_g+L_h+1)} \tag{18}$$

$$\tilde{\boldsymbol{H}} := \mathcal{H}|_{N=2L_g+L_h+1} \tag{19}$$

$$bC_{2x,d} := \tilde{\boldsymbol{C}}_{2x}(1:(L_g+1)P, dP+1:(L_g+1+d)P) \tag{20}$$

where we have used the MATLAB notation $\boldsymbol{A}(r_1:r_2, c_1:c_2)$ to indicate the submatrix of $\boldsymbol{A}$ consisting of rows r_1 through r_2 and columns c_1 through c_2. Note that $\boldsymbol{C}_{2x,d} = E\{\boldsymbol{x}_{L_g+1}(n)\boldsymbol{x}_{L_g+1}^{*'}(n-d)\}$, where the vector $\boldsymbol{x}_{L_g+1}(n)$ is defined similar to (4) with $y(n)$ replaced with $x(n)$. In addition, we note that

$$\tilde{\boldsymbol{H}}'(:, dP+1:(d+1+L_g)P) = \begin{bmatrix} \boldsymbol{0}_{dP\times(L_g+1)P} \\ \tilde{\boldsymbol{H}}'(1:P(L_h+L_g+1-d), 1:(L_g+1)P) \end{bmatrix} \tag{21}$$

where $\boldsymbol{0}_{Pd\times(L_g+1)P}$ is a $d\times(L_g+1)P$ matrix of zeros. Therefore [see (12)], we can write

$$\tilde{\boldsymbol{H}}'(:, dP+1:(d+1+L_g)P)\boldsymbol{g}_0 = \boldsymbol{e}_{d+1}. \tag{22}$$

From (22), we can write

$$\boldsymbol{C}_{2x,d}^{*}\boldsymbol{g}_0 = \sigma_w^2\tilde{\boldsymbol{H}}^{*}(1:(L_g+1)P,:) \cdot \tilde{\boldsymbol{H}}'(:, dP+1:(d+1+L_g)P)\boldsymbol{g}_0.$$

From (22), the above becomes

$$\begin{aligned}\boldsymbol{C}_{2x,d}^{*}\boldsymbol{g}_0 &= \sigma_w^2\tilde{\boldsymbol{H}}^{*}(1:(L_g+1)P,:)\boldsymbol{e}_{d+1} \\ &= \sigma_w^2\tilde{\boldsymbol{H}}^{*}(1:(L_g+1)P, d+1) \\ &= \sigma_w^2\boldsymbol{H}^{*}(:, d+1).\end{aligned} \tag{23}$$

Following the same development as for the zero-delay equalizer, the delay d equalizer satisfies

$$\boldsymbol{C}_{2x}^{*}\boldsymbol{g}_d = \sigma_w^2\boldsymbol{H}^{*}(:, d+1). \tag{24}$$

Equating (23) and (24) yields

$$\boldsymbol{C}_{2x}^{*}\boldsymbol{g}_d = \boldsymbol{C}_{2x,d}^{*}\boldsymbol{g}_0. \tag{25}$$

To solve for $\boldsymbol{g}_d$, we have [see (17)]

$$\boldsymbol{g}_d = \sigma_w^2[\boldsymbol{C}_{2x}^{*}]^{\dagger}\boldsymbol{C}_{2x,d}^{*}[\boldsymbol{C}_{2x}^{*}]^{\dagger}\boldsymbol{H}^{*}(:,1). \tag{26}$$

In other words, only one matrix (pseudo-) inverse is required to find $\boldsymbol{g}_d$ directly from the output correlation matrix.

The connection between the approach of (26) and the linear prediction approach of [29] provides an interesting topic for further research. We note, however, that for a delay of d, the method of [29] requires $\boldsymbol{h}(d) = [h_0(d) \;\; h_1(d) \;\; \cdots \;\; h_{P-1}(d)]$ to be known, and hence, the ambiguity (without estimating the channel *a priori*) is not a simple scale.

In many cases, it has been observed [26, ch. 10] that selecting $d \approx (L_g+L_h)/2$ (approximately the middle of the range of delays) results in good equalizer performance. Selection of the optimum delay d has been discussed in other papers (cf., [8] and [29]). Here, we show how, from the minimum delay $(d=0)$, ZF equalizers corresponding to any delay d can be obtained. In addition, [8] and [9] have proposed methods to find the ZF equalizers for all delays simultaneously, which, although it allows more design flexibility, is naturally more computationally intensive than the method proposed here.

At this point, we wish to note that while this paper addresses only i.i.d. inputs, this method has been extended to either colored or deterministic inputs [9], [10]. Blind equalization of colored or deterministic inputs is particularly useful for coded inputs and image processing [10].

Remark 2: The restriction to casual channels (shift ambiguity) and the limitation of finding $\boldsymbol{g}_d$ to within an unknown scale is not unique to this method. Rather, all blind identification and equalization methods are subject to these limitations (cf., [22], [28], [31]). In practice, these restrictions are usually overcome through the use of automatic gain/power control and differential encoding [26, ch. 12].

B. Blind MMSE Equalizers

While the previous section provided a blind method for finding an FIR ZF equalizer, the ZF equalizer does not address noise suppression. Hence, we are motivated to look for an alternative criterion for finding $\{g_i^{(d)}(n): i=0,1,\cdots,P-1\}_{n=0}^{L_g}$. Toward that goal, we consider the FIR Wiener filter, which provides the minimum mean-square sense linear estimate of $w(n-d)$ based only on $y(n)$ (blind MMSE).

Our goal is to find the $\{g_i^{(d)}(n): i=0,1,\cdots,P-1\}_{n=0}^{L_g}$ such that $J_d := E\{|\hat{w}(n)-w(n-d)|^2\}$ is minimized. As usual, we substitute (see Fig. 3)

$$\hat{w}(n) = \sum_{i=0}^{P-1}\sum_{\ell=0}^{L_g} g_i^{(d)}(\ell)y_i(n-\ell) \tag{27}$$

into J_d, take the complex derivative with respect to the unknown equalizer coefficients, and set these to zero

$$\frac{\partial}{\partial g_k^{(d)*}(m)}E\left\{\left|\sum_{i=0}^{P-1}\sum_{\ell=0}^{L_g} g_i^{(d)}(\ell)y_i(n-\ell) - w(n-d)\right|^2\right\} = 0$$

for $k=0,1,\cdots,P-1$ and $m=0,1,\cdots,L_g$. After simplification, this yields the orthogonality condition

$$\sum_{i=0}^{P-1}\sum_{\ell=0}^{L_g} g_i^{(d)}(\ell)\cdot E\{y_i(n-\ell)y_k^{*}(n-m)\} - E\{w(n-d)y_k^{*}(n-m)\} = 0. \tag{28}$$

Using the assumption that the noise is uncorrelated with input and that the input symbols are i.i.d., we have

$$\begin{aligned} &E\{w(n-d)y_k^*(n-m)\} \\ &\quad = E\left\{ w(n-d)\cdot \sum_{\ell} w^*(\ell)h_k^*(n-m-\ell)\right\} \\ &\quad = \sigma_w^2 h_k^*(d-m). \end{aligned} \tag{29}$$

Substitution of (29) into (28) yields

$$\sum_{i=0}^{P-1}\sum_{\ell=0}^{L_g} E\{y_i(n-\ell)y_k^*(n-m)\}g_i^{(d)}(\ell) = \sigma_w^2 h_k^*(d-m) \tag{30}$$

which can be written in matrix form for $m = 0, 1, \cdots, L_g$ and $k = 0, 1, \cdots, P-1$ as

$$E\{\boldsymbol{y}_{L_g+1}^*(n)\cdot \boldsymbol{y}_{L_g+1}'(n)\}\boldsymbol{g}_d = \sigma_w^2 \boldsymbol{H}^*(:, d+1). \tag{31}$$

Equation (31) then becomes [see (10)]

$$\boldsymbol{C}_{2y}^*\boldsymbol{g}_d = \sigma_w^2 \boldsymbol{H}^*(:, d+1). \tag{32}$$

This, however, is identical to (24) with $\boldsymbol{C}_{2x}$ (noise-free correlation) replaced by $\boldsymbol{C}_{2y}$ (signal plus noise correlation). Therefore, we can solve for the zero-delay MMSE equalizer and the delay d MMSE equalizer using the same procedures as in Section III-A. In particular, the nonzero delay $d \neq 0$ MMSE equalizer, by solving [see (23) and (26)]

$$\boldsymbol{g}_d = \sigma_w^2 [\boldsymbol{C}_{2y}]^{\dagger}\boldsymbol{C}_{2x,d}^*[\boldsymbol{C}_{2x}^*]^{\dagger}\boldsymbol{H}^*(:, 1). \tag{33}$$

In Section III-D, we briefly discuss selection of the delay d.

Remark 3: We note that (32) is valid for noise of any color. As with [22], (33) is valid for noise of any known color.

C. ZF—MMSE

In Section III-A, we presented an equalizer that was ZF and, hence, removed all of the ISI. However, ZF equalizers, unlike MMSE equalizers, do not address distortion from additive noise, and in some cases [26, p. 606], may unduly enhance the noise. Given that ZF equalizers are not always unique (see Theorem 1), it makes sense to search for the "best" ZF equalizer. In other words, we want to find the ZF equalizer that does the best job of suppressing the additive noise at the output of the equalizer.

From Fig. 3, at time n, the noise at the output of the equalizer $\epsilon(n) := \Sigma_i \Sigma_\ell g_i^{(d)}(\ell)v_i(n-\ell) = \boldsymbol{g}_d'\boldsymbol{v}_{Lg+1}(n)$ has variance

$$\sigma_\epsilon^2 := E\{\epsilon(n)\epsilon^*(n)\} = \boldsymbol{g}_d'\boldsymbol{C}_{2v}\boldsymbol{g}_d^* = \boldsymbol{g}_d^{*'}\boldsymbol{C}_{2v}^*\boldsymbol{g}_d. \tag{34}$$

Our goal is then to find the ZF FIR equalizer $\overline{\boldsymbol{g}}_d$ that minimizes σ_ϵ^2. Specifically, we seek the $\overline{\boldsymbol{g}}_d$ such that

$$\overline{\boldsymbol{g}}_d = \arg\min_{\boldsymbol{g}} \{\boldsymbol{g}_d^{*'}\boldsymbol{C}_{2v}^*\boldsymbol{g}_d\} \tag{35}$$

subject to the "statistical" ZF constraint [see (25)] $\boldsymbol{C}_{2x}^*\overline{\boldsymbol{g}}_d = \boldsymbol{C}_{2x,d}^*[\boldsymbol{C}_{2x}^*]^{-1}\sigma_w^2\boldsymbol{H}^*(:, 1)$. The constrained minimization problem can be solved using Lagrange multipliers for complex vector unknowns (e.g., [18, pp. 787–790]). The solution is (see Appendix A)

$$\begin{aligned} \overline{\boldsymbol{g}}_d = &[\boldsymbol{C}_{2v}^*]^{-1}\boldsymbol{C}_{2x}'\{\boldsymbol{C}_{2x}^*[\boldsymbol{C}_{2v}^*]^{-1}\boldsymbol{C}_{2x}'\}^{-1}\boldsymbol{C}_{2x,d}^*[\boldsymbol{C}_{2x}^*]^{-1} \\ &\cdot \sigma_w^2\boldsymbol{H}^*(:, 1). \end{aligned} \tag{36}$$

This equalizer is ZF (thus removing all the ISI in theory) and, among the class of ZFE's, does the best job of suppressing the noise gain. More precisely, for a given delay, it is the MMSE equalizer in the class of ZFE's. We note that when the additive noise is white, the solution of (36) becomes the minimum norm solution of (26). When the noise is colored, the minimum norm solution no longer does the "best" job of suppressing the noise at the output. This is illustrated in the simulations section (see Experiment 5).

Remark 4: For the SISO framework, [11] gives a nonparametric IIR version of the hybrid ZF-MMSE for cases where $h(n)$ is known or estimated *a priori*. Defining $\overline{G}_d(\omega)$ and $H(\omega)$ as the Fourier transforms of the IIR $\overline{g}_d(n)$ and FIR $h(n)$, respectively, the solution is (see Appendix B)

$$\begin{aligned} &\overline{G}_d\left(\frac{\omega-2\pi k}{P}\right) \\ &\quad = \frac{Pe^{-j\omega d}H^*\left(\dfrac{\omega-2\pi k}{P}\right)\Big/ S_{2v}\left(\dfrac{\omega-2\pi k}{P}\right)}{\displaystyle\sum_{\ell=0}^{P-1}\left|H\left(\frac{\omega-2\pi\ell}{P}\right)\right|^2\Big/ S_{2v}\left(\frac{\omega-2\pi\ell}{P}\right)} \end{aligned} \tag{37}$$

for $k = 0, 1, \cdots P-1, -\pi \leq \omega \leq \pi$, where $S_{2v}(\omega)$ is the power spectral density of the noise. This provides an IIR ZF equalizer, which minimizes the noise variance at the output. Since for stability and ease in implementation we prefer FIR equalizers, the IIR $\overline{G}_d(\omega)$ can be used to determine how fast the IIR $\overline{g}_d(n)$ decays to zero and then use this to design FIR $\boldsymbol{g}_d$ of appropriate length. A seemingly related IIR filter was reported in [2]. However, [2] provides an optimum IIR equalizer that minimizes the mean-square error but, contrary to (37), is not ZF.

D. Estimation from Sample Statistics

In practice, ensemble values are not available, and the matrix $\boldsymbol{C}_{2y}^*$ in (32) must be replaced by a consistent sample estimate based on the N-vector observations [24]

$$\hat{\boldsymbol{C}}_{2y}^* = \frac{1}{N}\sum_{\ell=0}^{N-1}\boldsymbol{y}_{L_g+1}^*(\ell)\boldsymbol{y}_{L_g+1}'(\ell) \tag{38}$$

where $\boldsymbol{y}_{L_g+1}(\ell)$ is defined as in (4).

Due to space limitations, a proof of the consistency for the equalizer estimates in Section III will not be shown. However,

we make the following comments on the proof. It is known that if the cumulants of the noise $v(n)$ are absolutely summable and the input has finite moments [4], then the normalized sample estimate $\hat{\boldsymbol{C}}^*_{2y}$ converges in the mean-square sense to the true value $\boldsymbol{C}^*_{2y}$, i.e.,

$$\frac{1}{N}\sum_{\ell=0}^{N-1}\boldsymbol{y}^*_{L_g+1}(\ell)\boldsymbol{y}'_{L_g+1}(\ell)\xrightarrow[N\to\infty]{\text{m.s.s.}}\boldsymbol{C}^*_{2y}.$$

In addition, the equalizer found by solving equations of the form of (16) or (25) will be unique, provided the conditions of Theorem 1 for uniqueness are satisfied. Under these conditions, (9) or (16) are one-to-one continuous functions of consistent estimators and, therefore, are consistent estimates of $\boldsymbol{g}_0$ and $\boldsymbol{g}_d$, respectively.

For the ZF and ZF-MMSE equalizers, the matrix $\hat{\boldsymbol{C}}_{2x}$ could be estimated by using $x(n)$ in place of $y(n)$ in (38). If the noise correlation $\boldsymbol{C}_{2v}$ is known or can be estimated from the noise-only data, then one may estimate the $\hat{\boldsymbol{C}}_{2y}$ as in (38) and find $\hat{\boldsymbol{C}}_{2x}$ according to

$$\hat{\boldsymbol{C}}_{2x}=\hat{\boldsymbol{C}}_{2y}-\boldsymbol{C}_{2v}. \tag{39}$$

We note that knowledge of $\boldsymbol{C}_{2v}$ is also required by the ZF methods of [8] and [28] and the channel identification method of [32]. When the additive noise is primarily due to thermal effects at the receiver, knowledge of $\boldsymbol{C}_{2v}$ is not unreasonable since the additive thermal noise is typically white, and any coloring will be due to known effects such as matched filtering.

For the MMSE case, d typically is chosen to be in the middle of the range of delays $[0, L_g+L_h]$. Alternatively, we can try to find the "best" delay in terms of smallest MSE, as we discuss next. The MMSE $J_d^{\min}$ can be evaluated using the orthogonality principle (28)

$$\begin{aligned}J_d^{\min}&=E\{-w^*(n-d)[\hat{w}(n)-w(n-d)]\}\\&=\sigma_w^2-E\{\hat{w}(n)w^*(n-d)\}\\&=\sigma_w^2-\sum_{i=0}^{P-1}\sum_{\ell=0}^{L_g}g_i^{(d)}(\ell)E\{y_i(n-\ell)w^*(n-d)\} \quad (40)\\&=\sigma_w^2-\sum_{i=0}^{P-1}\sum_{\ell=0}^{L_g}g_i^{(d)}(\ell)h_i(d-\ell) \quad (41)\end{aligned}$$

where in deriving the last equality, we used the whiteness of $w(n)$. The MMSE in (40) shows clearly its dependence on d and $\{h_i(n)\}_{i=0}^{P-1}$ and suggests the following brute-force algorithm for selecting the MMSE delay as

$$\hat{d}=\arg\min_d J_d^{\min}. \tag{42}$$

The resulting delay selection algorithm is summarized in the following steps:

Step 1: Solve (32) with $d=0$ to obtain $\boldsymbol{g}_0$, and solve (33) to obtain $\boldsymbol{g}_d$.

Step 2: Deconvolve to obtain $\hat{w}(n)$ using (27) with $d=0$.

Step 3: Estimate the cross-correlation in (40) using the sample average

$$\hat{h}_i(d-\ell)=\frac{1}{N}\sum_{n=0}^{N-1}y_i(n-\ell)\hat{w}^*(n-d). \tag{43}$$

Step 4: Evaluate $\hat{J}_d^{\min}$ via (41) for all $d=0,1,\cdots,L_h+L_g$, and select $\hat{d}$ as in (42).

A study of the MSE achieved by this blind MMSE equalizer in comparison with the results for the constant modulus algorithm (CMA) provides an interesting topic for future work. In addition, [7] purportedly contains details on how zero locations and delay affect the MMSE.

IV. Recursive/Adaptive Blind ZF Equalizers

The previous section describes batch methods for finding the equalizer taps from the vector output correlations without first explicitly finding the channel estimates. One major advantage "direct" equalizer methods have over the two-step procedures is the ability to develop computationally attractive recursive methods for estimating the equalizers. Similarly, direct methods allow for the development of adaptive implementations that are useful for tracking time-varying channels. In this section, we describe two adaptive (recursive) methods for finding the equalizer taps directly from the data. The first is a cyclic version of the recursive least-squares (RLS) algorithm, whereas the second is a stochastic gradient descent algorithm that we call the cyclic least mean square (LMS) algorithm because of the cyclo-multirate equivalence alluded to in (2).

A. RLS Equalizer

To update the correlation estimates recursively, we use the sample estimator

$$\begin{aligned}\hat{\boldsymbol{C}}^*_{2y}(T)&=\sum_{\ell=0}^{T-1}\lambda^{T-\ell}\boldsymbol{y}^*_{L_g+1}(\ell)\boldsymbol{y}'_{L_g+1}(\ell)\\&\quad+\boldsymbol{y}^*_{L_g+1}(T)\boldsymbol{y}'_{L_g+1}(T)\\&=\lambda\hat{\boldsymbol{C}}^*_{2y}(T-1)+\boldsymbol{y}^*_{L_g+1}(T)\boldsymbol{y}'_{L_g+1}(T) \quad (44)\end{aligned}$$

where $\hat{\boldsymbol{C}}_{2y}(T)$ and $\hat{\boldsymbol{C}}_{2y}(T-1)$ are the sample correlation matrices at time T and $T-1$, respectively, and $\boldsymbol{y}_{L_g+1}(m)$ is defined as in (4). The term λ $(0\le\lambda\le 1)$ is a "forgetting" factor included to reduce the influence of past observations on the statistics and thereby allow the correlation estimates to follow time variations in the channel. Note, however, that when $\lambda<1$, (44) is a biased estimate of the ensemble value $\boldsymbol{C}^*_{2y}$.

Defining $\boldsymbol{P}(T):=[\hat{\boldsymbol{C}}^*_{2y}(T)]^{-1}$ and using the matrix inversion lemma (e.g., [18, p. 480]) with (44), we can write

$$\boldsymbol{P}(T)=\lambda^{-1}\boldsymbol{P}(T-1)-\lambda^{-1}\boldsymbol{k}(T)\boldsymbol{y}'_{L_g+1}(T)\boldsymbol{P}(T-1) \tag{45}$$

and

$$\boldsymbol{k}(T):=\frac{\lambda^{-1}\boldsymbol{P}(T-1)\boldsymbol{y}^*_{L_g+1}(T)}{1+\lambda^{-1}\boldsymbol{y}'_{L_g+1}(T)\boldsymbol{P}(T-1)\boldsymbol{y}^*_{L_g+1}(T)}. \tag{46}$$

Notice that (45) and (46) do not require a matrix inverse and thereby provide a recursive method for computing $\boldsymbol{P}(T)$. From (32), we can write the equalizer vector estimate at time T as

$$\hat{\boldsymbol{g}}_0(T)=\boldsymbol{P}(T)\boldsymbol{f}(T). \tag{47}$$

The term $\boldsymbol{f}(T)$ is introduced for notational convenience and is defined, for $T = 0, 1, \cdots N-1$, as

$$\boldsymbol{f}(T) := \lambda \boldsymbol{f}(T-1) + \sigma_w^2 \boldsymbol{H}(:,1)$$

where $\boldsymbol{f}(-1) := 0$. This recursion naturally can be expressed in closed form by use of the geometric series

$$\boldsymbol{f}(T) = \frac{1-\lambda^{T+1}}{1-\lambda}\sigma_w^2 \boldsymbol{H}(:,1).$$

In summary, we have the following steps for computing $\hat{\boldsymbol{g}}_0(T)$ at each time T (see Remark 5 for comments on initialization).

Step 1:

$$\boldsymbol{k}(T) = \frac{\lambda^{-1}\boldsymbol{P}(T-1)\boldsymbol{y}^*_{L_g+1}(T)}{1+\lambda^{-1}\boldsymbol{y}'_{L_g+1}(T)\boldsymbol{P}(T-1)\boldsymbol{y}^*_{L_g+1}(T)}.$$

Step 2:

$$\boldsymbol{P}(T) = \lambda^{-1}\boldsymbol{P}(T-1) - \lambda^{-1}\boldsymbol{k}(T)\boldsymbol{y}'_{L_g+1}(T) \times \boldsymbol{P}(T-1).$$

Step 3:

$$\boldsymbol{f}(T) = \frac{1-\lambda^{T+1}}{1-\lambda}\sigma_w^2 \boldsymbol{H}(:,1).$$

Step 4:

$$\hat{\boldsymbol{g}}_0(T) = \boldsymbol{P}(T)\boldsymbol{f}(T).$$

The form of the estimate in (47) does not require a matrix inverse and, hence, is computationally feasible for adaptive implementation [c.f. (45)], or if $\lambda = 1$ is chosen, (47) provides a method to recursively compute the time-invariant equalizer taps, thereby reducing the memory requirements for long data records. In addition, while this procedure gives an adaptive estimate of the zero-delay MMSE equalizer, it is straightforward to modify to provide an adaptive estimate of either the delay d ZF or the d delay MMSE equalizer.

The forgetting factor λ, which provides for tracking of "slow variations" in the true correlations, affects the convergence of $\hat{\boldsymbol{C}}_{2y}(T)$ and, thus, the accuracy of the estimator. The tradeoff between tracking and convergence properties dictates the choice of λ usually chosen in the interval [0.98, 1].

Remark 5: Initialization of the RLS procedure is an important issue that, in a different context, has been studied by other researchers [18, ch. 13]. In many cases, it is reasonable to use a small set of observations $T_0 \ll N$, where N is the number of vector observations, to find an initial estimate of $\boldsymbol{P}(T_0)$ and $\boldsymbol{g}_0(T_0)$. Since the algorithm is initialized with batch estimates, the recursive estimates are exactly equal (when $\lambda = 1$) to their batch counterparts. When the channel is time invariant, the latter establishes the consistency of the RLS estimate because the batch method is consistent (see Section III-D). For $\lambda < 1$ or other initializations, convergence analysis is beyond the scope of this paper.

B. Cyclic LMS Equalizer

The cyclic LMS blind equalizer updates the equalizer estimate at each symbol through applications of the gradient descent algorithm, i.e.,

$$\hat{\boldsymbol{g}}_0(T) = \hat{\boldsymbol{g}}_0(T-1) - \tfrac{1}{2}\mu \nabla \hat{J}_0(T) \tag{48}$$

where $\hat{\boldsymbol{g}}_0(T)$ and $\hat{\boldsymbol{g}}_0(T-1)$ are the equalizer vectors at time T and $T-1$, respectively, μ is the stepsize, and $\nabla \hat{J}_0(T)$ is the instantaneous approximation at time T to the gradient of the cost function $J_0 := E\{|\hat{w}(n) - w(n)|^2\}$, which is

$$\nabla J_0 = \left[\frac{\partial J_0}{\partial g_0^*(0)} \frac{\partial J_0}{\partial g_1^*(0)} \cdots \frac{\partial J_0}{\partial g_{P-1}^*(L_g)}\right]'.$$

Following the same steps as in the derivation of the MMSE equalizer in Section III-B, we see

$$\nabla J_0 = E\{\boldsymbol{y}^*_{L_g+1} \quad \boldsymbol{y}'_{L_g+1}\}\boldsymbol{g}_0 - \sigma_w^2 \boldsymbol{H}^*(:,1). \tag{49}$$

The instantaneous approximation at time T is obtained by

$$\nabla \hat{J}_0(T) := \boldsymbol{y}^*_{L_g+1}(T)\boldsymbol{y}'_{L_g+1}(T)\hat{\boldsymbol{g}}_0(T-1) - \sigma_w^2 \boldsymbol{H}^*(:,1). \tag{50}$$

As with any gradient search method, initialization and choice of stepsize μ play critical roles in the speed of convergence and the steady-state performance. These issues and convergence analysis have been studied by many researchers in other contexts (e.g., [18, ch. 9]). For FSE's, such a study is beyond the scope of this paper.

Despite the potential for steady-state misadjustment error and slow convergence, the cyclic LMS has extremely low computational complexity. In addition, it does not rely on an explicit matrix inverse, and therefore, as we demonstrate in Experiment 3, is not as sensitive to nearly common subchannel roots. When subchannel roots are close to being common, the channel matrix $\boldsymbol{H}$ becomes ill conditioned. Correspondingly, methods depending on rank properties of $\boldsymbol{H}$ or $\boldsymbol{C}_{2x}$ [see (14)], such as [22], [28], or [31], yield estimates with larger variance. In contrast, the affect on the LMS algorithm is to "slow" the convergence [18, pp. 334–335].

V. Simulation Results—Comparisons using Real Data

In this section, we use simulations to examine the performance of the direct equalization methods described in Section III and the adaptive/recursive methods described in Section IV. In addition, we compare the performance of the proposed methods with existing second- and higher order methods for equalization using an empirically measured digital microwave channel. As performance measures, we estimate the mean-square symbol error, MSE $:= E\{|\hat{w}(n-d) - w(n)|^2\}$ and the residual ISI (cf., [27]) over 100 Monte Carlo runs. If $\{\hat{g}_i^{(m)}(n): i = 0, 1, \cdots P-1\}_{n=0}^{L_g}$ is the equalizer estimate obtained on the mth Monte Carlo run, the residual ISI is defined as

$$\text{ISI}\,(m) := \frac{\sum_n |f^{(m)}(n)|^2 - \max_n |f^{(m)}(n)|^2}{\max_n |f^{(m)}(n)|^2} \tag{51}$$

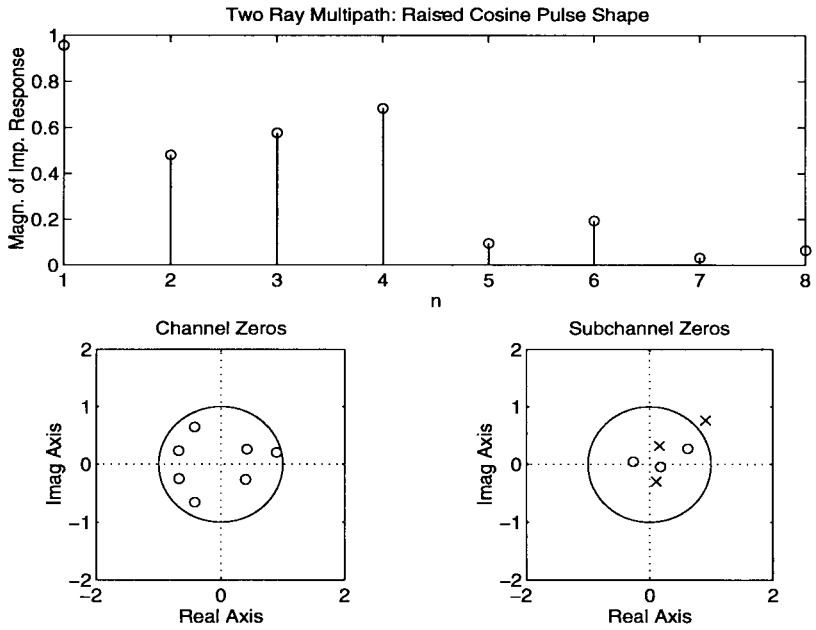

Fig. 4. Two-ray multipath channel: Experiment 1.

where $f^{(m)}(n)$ is the "overall" channel impulse response

$$f^{(m)}(n) := \sum_{i=0}^{P-1} \sum_{\ell=0}^{L_g} \hat{g}_i^{(m)}(\ell) h_i(n-\ell). \tag{52}$$

To find an expression for mean-square symbol error, we first substitute $\hat{w}(n-d) = \Sigma_\ell\, w(\ell) f^{(m)}(n-\ell) + \Sigma_n\, v(n) g(n-\ell)$ into the MSE definition and then simplify using the whiteness of $w(n)$. On the mth Monte Carlo run, the MSE is then

$$\begin{aligned} &\mathrm{MSE}\,(m) \\ &\quad := \sigma_w^2 \left[\sum_n |f^{(m)}(n)|^2 - f^{(m)}(d) - f^{*(m)}(d) + 1 \right] \\ &\qquad + \sum_{k_1} \sum_{k_2} c_{2v}(k_2 - k_1) \hat{g}^{(m)}([n-d]P - k_1) \\ &\qquad \cdot \hat{g}^{*(m)}([n-d]P - k_2) \end{aligned} \tag{53}$$

where d is the desired delay. For all simulations, we have defined the signal-to-noise ratio (SNR) to be at the input to the equalizer, or

$$\mathrm{SNR} := \frac{E\{|x(n)|^2\}}{E\{|v(n)|^2\}}.$$

For each experiment, we have used an i.i.d. input sequence drawn from a 16-QAM constellation and have assumed the oversample ratio to be 2 (i.e., $P = 2$). For Experiments 1–4, the noise is drawn from a white Gaussian distribution at varying SNR's. A receive filter with bandwidth P/T_s guarantees that $v(n)$ is white.

Experiment 1—Performance in Noise: We first consider the performance of the MMSE equalizer (Section III-B) in the presence of additive noise $v(n)$. The channel is a causal approximation to a two-ray multipath mobile radio environment. The continuous-time channel spans four symbols and is described for $t \in [0, 4T)$ by

$$h_c(t) = e^{-j2\pi(0.15)} r_c(t - 0.25T, \beta) + 0.8 e^{-j2\pi(0.6)} r_c(t - T, \beta)$$

where $r_c(t, \beta)$ is the raised-cosine [26, p. 546] with roll-off factor β. For this experiment, $\beta = 0.35$. The discrete-time equivalent channel is found by sampling $h_c(t)$ at a rate of $T/2$ or $h(n) := h_c(nT/2)$ for $n = 0, 1, \cdots, 7$. Fig. 4 shows the magnitude of the impulse response, the zeros of $h(n)$, and the zeros of the subchannels $h_0(n)$ and $h_1(n)$. Fig. 5 depicts the mean-square symbol error for a zero-delay equalizer whose subchannels are of order $L_g = 3$ across 2500 symbols for several SNR's. Fig. 6 shows the residual ISI for the same equalizers. Figs. 5 and 6 show that the FSE does a good job of equalization after as few as 500 symbols. Table I compares the performance of the FSE after 2000 symbols with the optimum Wiener filter, which has exact knowledge of the channel (cf., [6]). We see that the blind MMSE is coming close to an ideal linear equalizer in performance. Finally, Fig. 7 shows the received constellation and the equalized constellation at SNR =25 dB for 750 symbols. Clearly, the blind equalizer has opened the eye.

TABLE I
COMPARISON OF BLIND FSE AND OPTIMUM LINEAR EQUALIZER

SNR in dB	MSE	
	Blind FSE	Optimum MMSE
30	0.0066	0.0033
20	0.0274	0.0237
15	0.0652	0.0618
12.5	0.1026	0.0993

Experiment 2—Comparison with Existing Algorithms: In this experiment, we compare the performance of the recursive algorithms in Section IV with existing blind FS equalization techniques. The channel used is an empirically measured $P = 2$ digital microwave channel with duration spanning eight

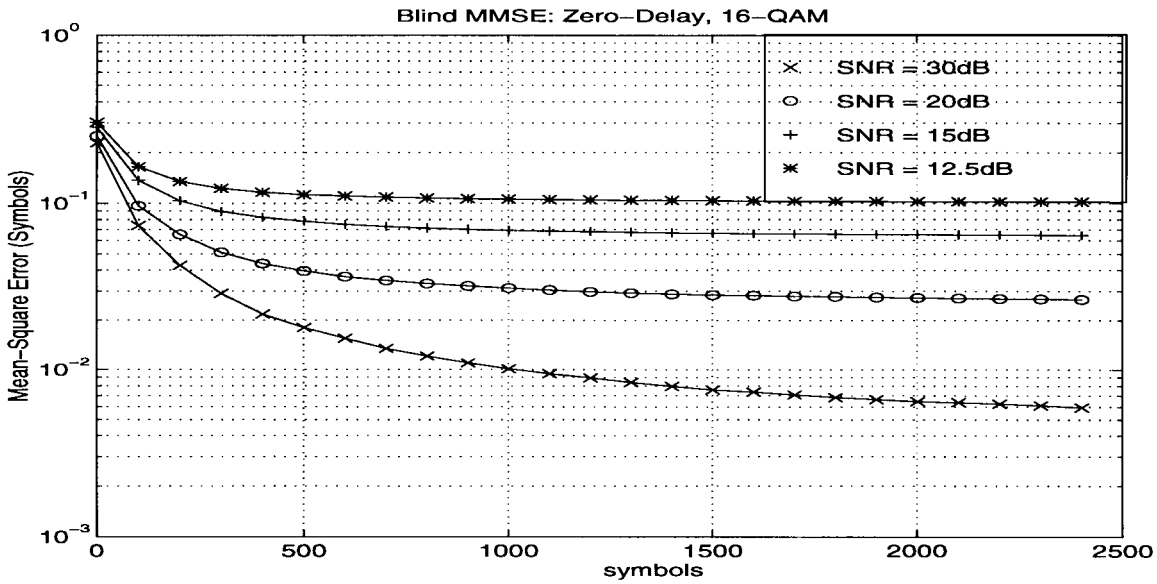

Fig. 5. Mean-square symbol error: Direct MMSE.

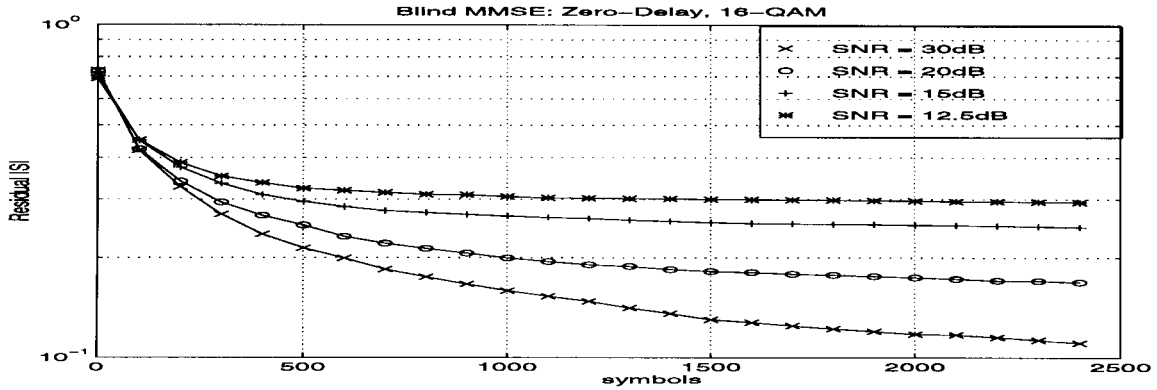

Fig. 6. Residual ISI: Direct MMSE.

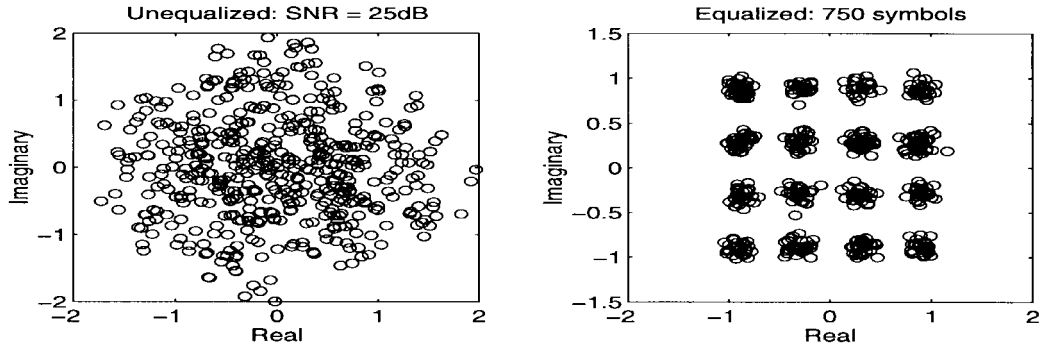

Fig. 7. Eye diagram: Direct MMSE.

symbols (see [6] for details on the channel). The SNR was 30 dB, and the equalizer is causal and of order $L_g = 7$. Based on this brute-force method, we selected $d = 4$ for the cyclic RLS and cyclic LMS algorithm. Fig. 8 shows the MSE for the cyclic RLS (Section IV-A), the cyclic LMS (Section IV-B), the CMA of the type $p = 2$ (CMA 2-2) [14], and the indirect (channel estimation followed by Wiener filter) methods of [22] and [32].

Trading off speed of convergence with steady-state error, the performance of both the cyclic LMS and the CMA depends on the stepsize, which in this experiment was chosen to be $\mu = 2.5 \times 10^{-3}$ for both. For the cyclic LMS algorithm and the cyclic RLS algorithm, 100 symbols were used to compute an initial estimate of the equalizer. Since the noise correlation is rarely known and is required by (33), the cyclic RLS and LMS algorithms used for Fig. 8 do not assume knowledge of

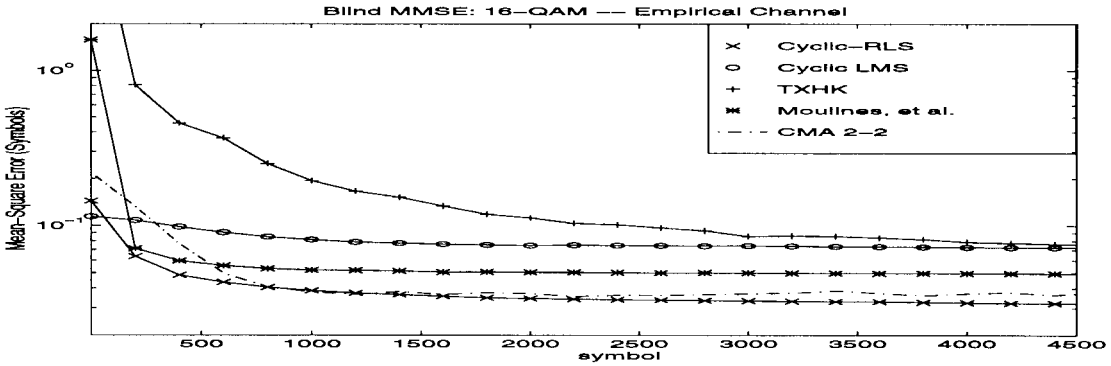

Fig. 8. Mean-square symbol error: Comparison at SNR = 30 dB.

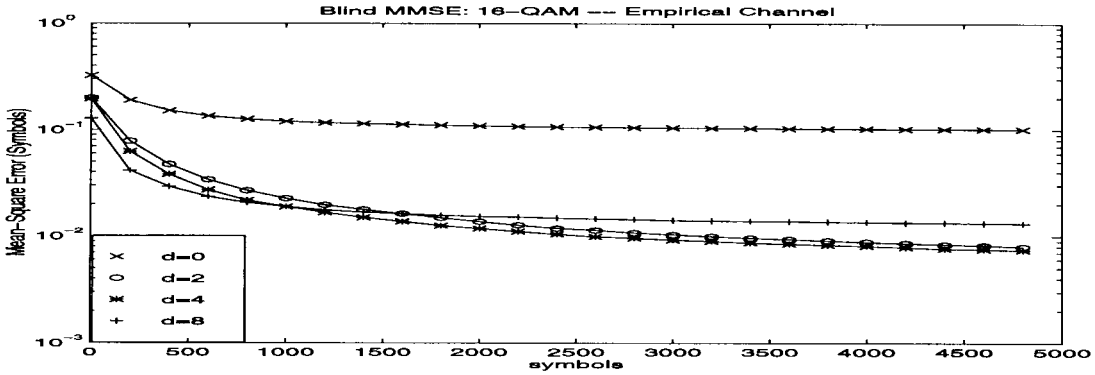

Fig. 9. Comparison of MSE for delays d: Cyclic RLS at SNR = 30 dB.

the noise-free correlation. Rather, they use $\boldsymbol{C}_{2y}$ in place of $\boldsymbol{C}_{2x}$ in (33). For the indirect methods, a new channel estimate and a new Wiener filter estimate are calculated every 400 samples. For the method of [32], we assume the algorithm has exact knowledge of channel order as well as knowledge of the noise power. For the method of [22], we chose a window size of 16 and assumed the algorithm had exact knowledge of the channel order. From Fig. 8, we see that the cyclic RLS algorithm has the lowest MSE and the fastest "convergence," although CMA 2-2 has similar large sample MSE. The Moulines *et al*. [22] method also shows fast MSE convergence. However, for each 400 samples, the algorithm of [22] requires two eigendecompositions to estimate the channel followed by a matrix inverse to compute the equalizer. Therefore, [22] is the most computationally intensive algorithm considered in Experiment 2.

MSE performance differences between cyclic RLS and [22] are channel dependent and may not be as pronounced as in Fig. 8. However, at least for this empirically measured channel, cyclic RLS outperformed [22]. The cyclic LMS outperforms also the indirect method of Tong *et al.* [32] and is computationally simpler than the cyclic RLS or the algorithm of [22].

Fig. 9 shows the MSE across a range of delays for the batch MMSE algorithm (33) and demonstrates the need for estimating the $d>0$ equalizer. For Fig. 9, we assumed the RLS algorithm has knowledge of the noise correlation.

TABLE II
NEARLY COMMON SUBCHANNEL ROOTS

Subchannel Roots	
$h_0(n)$	$h_1(n)$
0.0074 - 0.4999i	0.0070 - 0.5094i
-0.2736 - 0.3719i	-0.2777 - 0.3799i
-0.4560 + 0.2706i	-0.4383 + 0.2598
-0.2472 + 0.3870i	-0.2328 + 0.3628i
0.5213 - 0.0539i	0.4902 - 0.0545i
0.1961 + 0.7765i	0.1008 + 0.2539i
0.4149 + 0.1422i	0.3789 + 0.1586i

Experiment 3—Comparison (Near Common Root Channel): It is known that blind equalization methods based on the second-order statistics of fractionally sampled channels do not perform well when the subchannels are close to being common (cf., [6]). In this experiment, we compare the performance of the cyclic LMS (Section IV-B), the CMA 2-2, and the indirect method of Moulines *et al*. [22] using the same parameters as in Experiment 2. The channel used for this experiment however has 6 roots that are "close" to being common. Table II gives the root locations, whereas Fig. 10 shows graphically the magnitude of the channel impulse response, the channel zeros, and the subchannel zeros. Figs. 11 and 12 show the MSE and the ISI, respectively.

As we see from Figs. 11 and 12, the cyclic LMS (being a criterion-based approach) shows robustness to this nearly common root channel. Subspace (and cyclic RLS) methods

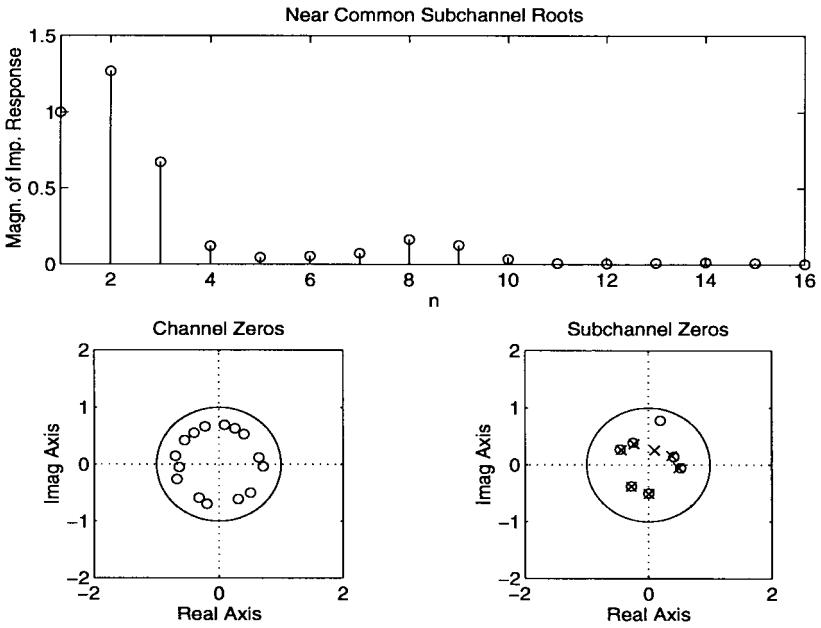

Fig. 10. Near-common roots channel: experiment 3.

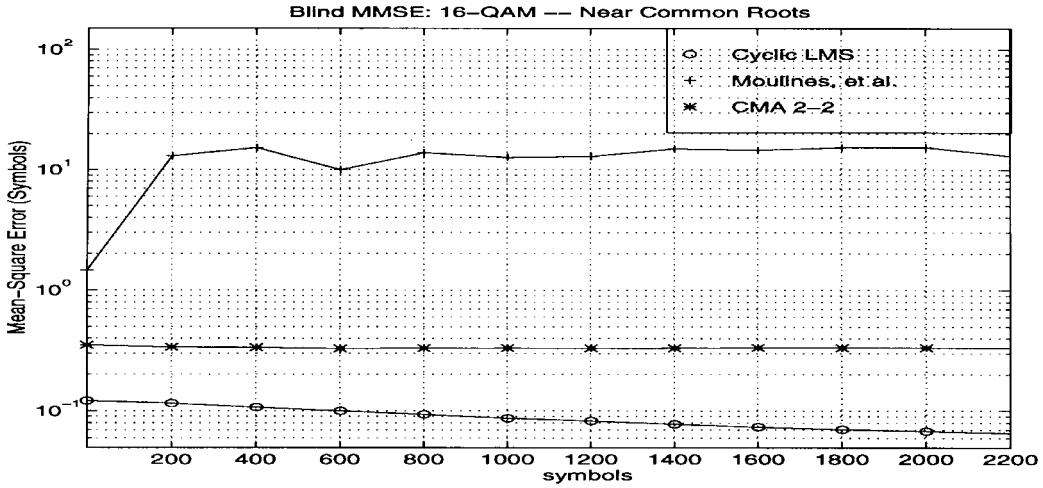

Fig. 11. Mean-square symbol error: Comparison at SNR = 30 dB.

perform batch (or recursive) inversion of ill-conditioned correlation matrices in this case. The price paid by LMS is slower convergence when channels have near common roots.

Experiment 4—Time-Varying Channels: In this experiment, we investigate the performance of the adaptive procedures of Section IV for a time-varying channel. As in Experiment 1, the channel is a causal approximation to a two-ray multipath channel with duration spanning four symbols. The continuous-time channel $h_c(t)$ for $t \in [0, 4T)$ is described by

$$h_c(t) = \sum_{\ell=0}^{1} a_\ell(t) r_c(t - \gamma_\ell, \beta) \tag{54}$$

where the roll-off factor β is again set to 0.35, $\gamma_0 = 0.25$, and $\gamma_1 = 1$. The discrete-time channel $h(n)$ is obtained by $h(n) = h_c(nT/2)$ for $n = 0, 1, \cdots, 7$. For Experiment 4, the "gain" of each path changes with time. Table III shows the values for $a_\ell(t) = |a_\ell(t)| \exp\{-j \angle a_\ell(t)\}$. We note that

TABLE III
TIME-VARYING CHANNEL: COEFFICIENTS

Time t	$a_0(t)$		$a_1(t)$	
	Magn. $\lvert a_0(t)\rvert$	Phase $\angle a_0(t)$	Magn. $\lvert a_1(t)\rvert$	Phase $\angle a_1(t)$
$t < 2{,}250$	1	2π 0.15	0.8	2π 0.6
$t \geq 2{,}250$	1	2π 0.45	0.4	2π 0.1

for $t < 2,250$, this is the same channel as in Experiment 1. Fig. 13 shows the MSE for this time-varying channel where $L_g = 3, \lambda = 0.9980$ (for cyclic RLS), $\mu = 2.5 \times 10^{-3}$ (for cyclic LMS), and the SNR = 20 dB. From Fig. 13, we see that both algorithms are able to adapt and equalize the time-varying channel. Fig. 14 shows the estimated MSE of the direct equalizer across a range of delays [see (41)].

Experiment 5—Hybrid ZF-MMSE Performance: In this experiment, we compare the performance of the zero-delay ZF

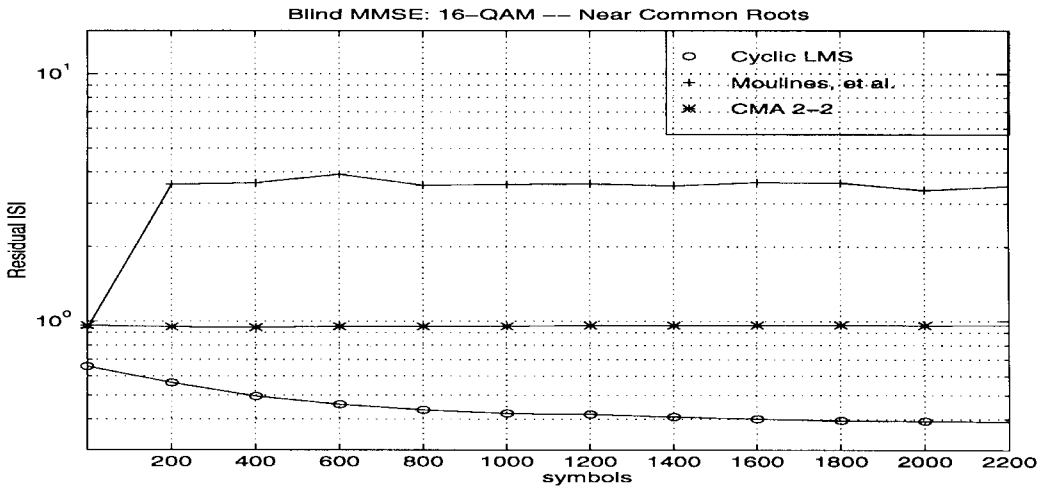

Fig. 12. Residual ISI: Comparison at SNR = 30 dB.

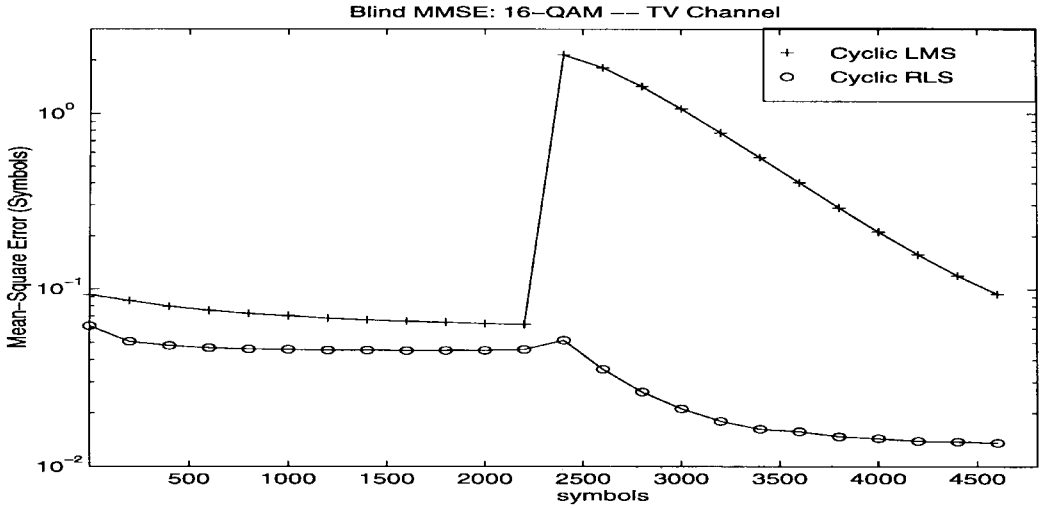

Fig. 13. Mean-square symbol error: Time-varying channel at SNR = 20 dB.

equalizer (Section III-A), the blind MMSE (Section III-B), and the hybrid ZF-MMSE (Section III-C) for the channel of Experiment 1. For this experiment, the additive noise $v(n)$ is no longer white but is colored and yields an SNR $= 5$ dB. The additive noise is generated by passing white Gaussian noise through a moving average process with coefficients 1, $0.5 - 0.25$j, $0.5 - 0.1$j, and $0.4 + 0.3$j. Table IV shows the ISI, the MSE, the average equalizer norm, and the noise power after 1000 symbols. The noise power σ_ϵ^2, which was defined in (17), is the gain of the noise induced by the equalizer. The norm of the equalizer on the mth Monte Carlo run is given by $|\hat{\boldsymbol{g}}_0^{(m)}|^2 := \Sigma_{i=0}^{1} \Sigma_{n=0}^{7} |g_i(n)|^2$. The equalizer was order 7 $(L_g = 7)$ and, therefore (see Theorem 1), was not unique. From Table IV, we see that each equalizer is indeed minimum with respect to its design criterion. The ZF and the ZF-MMSE both achieve the same residual ISI, which is much smaller than that achieved by the MMSE equalizer. The difference between the ZF and the ZF-MMSE can be seen by comparing the noise power and the equalizer norm. The ZF-MMSE achieves smaller noise power (and therefore MSE) but is not the minimum norm solution (obtained by solving (17) using the pseudo-inverse) given by the ZF equalizer. As expected, the MMSE equalizer has the smallest MSE.

VI. Conclusions and Future Directions

Using output-only data, we have provided three procedures for directly calculating FS FIR equalizers based on second-order statistics. These procedures do not require channel estimation as a first step, and each one possesses a different optimality. One method eliminates all the ISI or is zero-forcing, another is minimum mean-square error, and the last is a hybrid method that provides the minimum mean-square error zero-forcing equalizer.

To cope with slowly varying time-varying channels and/or provide computationally efficient implementations, we developed two adaptive methods for finding the time-varying equalizer taps directly from the data. One method is based on recursive least squares, and the other on a stochastic gradient descent or LMS-like algorithm. The latter is computationally simpler than the former, and both exhibit "robustness" to ill conditioning caused by near-common subchannel zeros

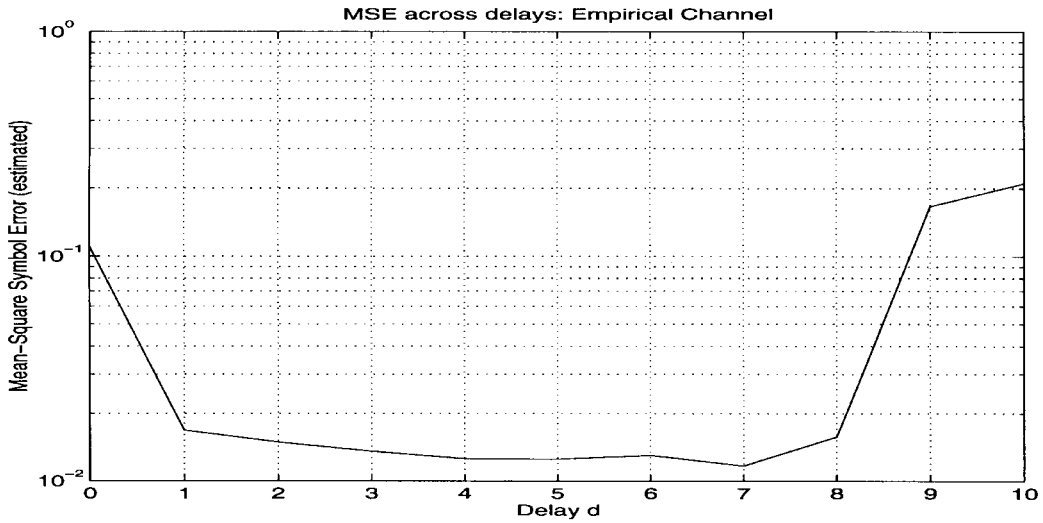

Fig. 14. Selection of optimum delay: Empirical channel.

TABLE IV
COMPARISON OF DIRECT EQUALIZERS FOR COLORED NOISE

	ISI	MSE	Noise Power	Eq. Norm
ZF	0.2741	4.0519	4.0417	12.7446
MMSE	0.6464	0.5081	0.2296	0.7237
ZF-MMSE	0.2741	4.0402	4.030	12.7815

because they avoid matrix inversion that is present with existing methods.

A preliminary comparison between the second- and higher order algorithms is given by [6], whereas a more comprehensive comparison is the topic of a future paper. In addition, convergence and initialization issues for the two adaptive procedures will be studied. Finally, performance of weighted least squares versions [16] as well as links of the present approach with the linear prediction approach [8], [28], [29] are worth further investigation.

APPENDIX A
DERIVATION OF HYBRID FIR ZF-MMSE

We seek the equalizer that satisfies

$$\bar{\boldsymbol{g}}_d = \arg\min_{\boldsymbol{g}_d} \{\boldsymbol{g}_d^{*'} \boldsymbol{C}_{2v}^* \boldsymbol{g}_d\} \tag{55}$$

subject to the statistical ZF requirement

$$\boldsymbol{C}_{2x}^* \bar{\boldsymbol{g}}_d = \tilde{\boldsymbol{C}}_{2x,e}^* [\boldsymbol{C}_{2x}^*]^{-1} \sigma_w^2 \boldsymbol{H}(:,1).$$

This can be solved using the Lagrange multiplier approach [18, pp. 787–790]. First, we form the objective function

$$J(\boldsymbol{g}_d) = \boldsymbol{g}_d^{*'} \boldsymbol{C}_{2v}^* \boldsymbol{g}_d + \mathrm{Re}\{\boldsymbol{\lambda}^{*'} \cdot [\boldsymbol{C}_{2x} \boldsymbol{g}_d - \boldsymbol{C}_{2x,e}^* [\boldsymbol{C}_{2x}^*]^{-1} \cdot \sigma_w^2 \boldsymbol{H}(:,1)]\} \tag{56}$$

where $\boldsymbol{\lambda}$ is the vector of complex Lagrange multipliers. Now, we form the adjoint by taking the gradient with respect to $\boldsymbol{g}_d$.

$$\nabla_{\boldsymbol{g}_d} J(\boldsymbol{g}_d) = \boldsymbol{C}_{2v}^* \boldsymbol{g}_d + \tfrac{1}{2} \boldsymbol{C}_{2x}' \boldsymbol{\lambda}. \tag{57}$$

Setting (57) to zero and solving gives

$$\boldsymbol{g}_d = [\boldsymbol{C}_{2v}^*]^{-1} \cdot \boldsymbol{C}_{2x}' \cdot (-\tfrac{1}{2} \boldsymbol{\lambda}). \tag{58}$$

To solve for $-(1/2)\boldsymbol{\lambda}$, we substitute (58) into (25)

$$\boldsymbol{C}_{2x}^* [\boldsymbol{C}_{2v}^*]^{-1} \cdot \boldsymbol{C}_{2x}' \cdot (-\tfrac{1}{2} \boldsymbol{\lambda}) = \tilde{\boldsymbol{C}}_{2x,e}^* [\boldsymbol{C}_{2x}^*]^{-1} \sigma_w^2 \boldsymbol{H}(:,1).$$

Solving gives

$$-\tfrac{1}{2} \boldsymbol{\lambda} = [\boldsymbol{C}_{2x}^* [\boldsymbol{C}_{2v}^*]^{-1} \boldsymbol{C}_{2x}']^{-1} \tilde{\boldsymbol{C}}_{2x,e} [\boldsymbol{C}_{2x}^*]^{-1} \sigma_w^2 \boldsymbol{H}(:,1)$$

which when substituted into (58) gives (36).

APPENDIX B
DERIVATION OF IIR ZF-MMSE

We give a brief derivation of the IIR ZF-MMSE equalizer. In the SISO framework, the ZF condition is $\Sigma_m g(m) h(nP - m) = \delta(n - d)$. In the frequency domain, this becomes (cf., [23, p. 103])

$$\frac{1}{P} \sum_{\ell=0}^{P-1} G\left(\frac{\omega - 2\pi\ell}{P}\right) H\left(\frac{\omega - 2\pi\ell}{P}\right) = e^{-j\omega d}. \tag{59}$$

For a fixed ω, (61) can be written in vector form as

$$\boldsymbol{h}'(\omega) \cdot \boldsymbol{g}(\omega) - P e^{-j\omega d} = 0$$

where

$$\boldsymbol{h}(\omega) := \left[H\left(\frac{\omega}{P}\right) H\left(\frac{\omega - 2\pi}{P}\right) \cdots H\left(\frac{\omega - 2\pi(P-1)}{P}\right)\right]'$$

$$\boldsymbol{g}(\omega) = \left[G\left(\frac{\omega}{P}\right) G\left(\frac{\omega - 2\pi}{P}\right) \cdots G\left(\frac{\omega - 2\pi(P-1)}{P}\right)\right]'.$$

Similarly, the power spectral density of the filtered noise $\{\epsilon(n)\}_{n=0}^{NP-1}$ is

$$S_{2\epsilon}(\omega) := \lim_{T \to \infty} \frac{1}{NP} E\{|\mathcal{E}(\omega)|^2\} = \boldsymbol{g}^{*'}(\omega) \boldsymbol{S}_{2v}(\omega) \boldsymbol{g}(\omega)$$

where $\mathcal{E}(\omega)$ is the Fourier transform of $\epsilon(n)$, NP is the data length, and $\boldsymbol{S}_{2v}(\omega)$ is the diagonal matrix given by

$$\boldsymbol{S}_{2v}(\omega) := \operatorname{diag}\left[S_{2v}\left(\frac{\omega}{P}\right) \quad S_{2v}\left(\frac{\omega-2\pi}{P}\right) \quad \cdots \quad S_{2v}\left(\frac{\omega-2\pi(P-1)}{P}\right)\right].$$

Similar to Appendix A, we have the following minimization problem. We seek, at each frequency ω, the optimum IIR ZF equalizer that satisfies

$$\overline{\boldsymbol{g}}_d(\omega) = \arg\min_{\boldsymbol{g}_d} \frac{1}{P}\boldsymbol{g}_d^{*'}(\omega) \quad \boldsymbol{S}_{2v}(\omega) \quad \boldsymbol{g}_d(\omega) \tag{60}$$

subject to $\boldsymbol{h}'(\omega)\boldsymbol{g}(\omega) - P\exp\{-j\omega d\} = 0$. Using the Lagrange multiplier approach [18, pp. 787–790], we form the objective function

$$J(\boldsymbol{g}_d(\omega)) = \boldsymbol{g}_d^{*'}(\omega)\boldsymbol{S}_{2v}(\omega)\boldsymbol{g}_d(\omega) + \operatorname{Re}\{\lambda^*[\boldsymbol{h}'(\omega)\boldsymbol{g}_d(\omega) - Pe^{-j\omega d}]\}. \tag{61}$$

Taking the gradient with respect to $\boldsymbol{g}_d(\omega)$, setting to zero, and solving for $\overline{\boldsymbol{g}}_d(\omega)$ gives

$$\overline{\boldsymbol{g}}_d(\omega) = [\boldsymbol{S}_{2v}(\omega)]^{-1}\boldsymbol{h}^*(\omega)\frac{\lambda}{2}. \tag{62}$$

It is convenient to look at the individual entries of $\overline{\boldsymbol{g}}_d(\omega)$ in (62), e.g.,

$$G\left(\frac{\omega-2\pi k}{P}\right) = \frac{1}{2}\frac{1}{S_{2v}\left(\frac{\omega-2\pi k}{P}\right)}H^*\left(\frac{\omega-2\pi k}{P}\right)\lambda \tag{63}$$

for a given ω and for $k = 0, 1, \cdots, P-1$. Using (63), we find λ by substitution of (62) into the ZF constraint to obtain $\boldsymbol{h}'(\omega)[\boldsymbol{S}_{2v}(\omega)]^{-1}\boldsymbol{h}^*(\omega)\lambda/2 = P\exp\{-j\omega d\}$ or in scalar form

$$\lambda\sum_{\ell=0}^{P-1} H\left(\frac{\omega-2\pi\ell}{P}\right)H^*\left(\frac{\omega-2\pi\ell}{P}\right)\frac{1}{S_{2v}\left(\frac{\omega-2\pi\ell}{P}\right)} = 2Pe^{-j\omega d}. \tag{64}$$

This gives

$$\lambda = \frac{2Pe^{-j\omega d}}{\sum_{\ell=0}^{P-1}\left|H\left(\frac{\omega-2\pi\ell}{P}\right)\right|^2 \Big/ S_{2v}\left(\frac{\omega-2\pi\ell}{P}\right)}. \tag{65}$$

Finally, substitution of (65) into (63) yields the desired result.

ACKNOWLEDGMENT

The authors wish to thank the reviewers, in particular Reviewer B, for their helpful suggestions, and Dr. J. Treichler for providing the data on the empirically measured digital microwave channel. The second author wishes to thank T. Endres for his suggestions, help with simulations, and informative discussions on this topic.

REFERENCES

[1] K. Abed Meraim *et al.*, "Prediction error methods for time-domain blind identification of multichannel FIR filters," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, vol. 3, Detroit, MI, 1995, pp. 1968–1971.
[2] P. Balaban and J. Salz, "Optimum diversity combining and equalization in digital data transmission with applications to cellular mobile radio—Part I: Theoretical considerations," *IEEE Trans. Commun.*, vol. 40, pp. 885–894, May 1992.
[3] A. Benveniste and M. Goursat, "Blind equalizers," *IEEE Trans. Commun.*, vol. COMM-32, pp. 871–883, 1984.
[4] A. V. Dandawaté and G. B. Giannakis, "Asymptotic theory of mixed time averages and kth-order cyclic-moment and cumulant statistics," *IEEE Trans. Inform. Theory*, vol. 41, pp. 216–239, Jan. 1995.
[5] Z. Ding and Y. Li, "On channel identification based on second-order cyclic spectra," *IEEE Trans. Signal Processing*, vol. 42, pp. 1260–1264, May 1994.
[6] T. J. Endres, S. D. Halford, C. R. Johnson, Jr., and G. B. Giannakis, "Blind adaptive channel equalization using fractionally-spaced receivers: A comparison study," in *Proc. 30th Conf. Inform. Sci. Syst.*, Princeton Univ., Princeton, NJ, Mar. 1996.
[7] I. Fijalkow, A. Touzni, and C. R. Johnson, Jr., "Spatio-temporal equalizability under channel noise and loss of disparity," in *Proc. GRETSI*, 1995.
[8] D. Gesbert, P. Duhamel, and S. Mayrargue, "Subspace-based adaptive algorithms for the blind equalization of multichannel FIR filters," in *Proc. EUSIPCO*, 1994.
[9] G. B. Giannakis, "Direct blind equalizers of multiple FIR channels: A deterministic approach," submitted for publication.
[10] G. B. Giannakis and W. Chen, "Blind blur identification and multichannel image restoration using cyclostationarity," in *Proc. IEEE Workshop Nonlinear Signal Image Processing*, vol. II, Halkidiki, Greece, June 20–22, 1995, pp. 543–546.
[11] G. B. Giannakis and S. D. Halford, "Blind fractionally-spaced equalization of noisy FIR channels: Adaptive and optimal solutions," in *Proc. Intl. Conf. Acoust., Speech, Signal Processing*, vol. 3, Detroit, MI, 1995, pp. 1972–1975.
[12] G. B. Giannakis and E. Serpedin, "Linear multichannel blind equalizers of nonlinear FIR Volterra channels," in *Proc. 30th Conf. Inform. Sci. Syst.*, Princeton Univ., Princeton, NJ, Mar. 1996, pp. 1153–1158.
[13] G. B. Giannakis and M. K. Tsatsanis, "Restoring identifiability of fractionally-sampled blind channel estimators using HOS," in *Proc. Int. Conf. Higher Order Statist.*, Barcelona, Spain, June 1995, pp. 429–431.
[14] D. N. Godard, "Self-recovering equalization and carrier tracking in two dimensional data communication systems," *IEEE Trans. Commun.*, vol. COMM-28, pp. 1867–1875, Nov. 1980.
[15] S. D. Halford and G. B. Giannakis, "Channel order determination based on cyclic correlations," in *Proc. 28th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, 1994, pp. 425–429.
[16] ———, "Optimal blind equalization and symbol error analysis of fractionall sampled channels," in *Proc. 29th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Oct. 29–Nov. 1, 1995.
[17] D. Hatzinakos, "Nonminimum phase channel deconvolution using the complex cepstrum of the cyclic autocorrelation," *IEEE Trans. Signal Processing*, vol. 42, pp. 3026–3042, Nov. 1994.
[18] S. Haykin, *Adaptive Filter Theory*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1991.
[19] Y. Hua, "Fast maximum likelihood for blind identification of multiple FIR channels," *IEEE Trans. Signal Processing*, vol. 44, pp. 661–672, Mar. 1996.
[20] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1980.
[21] H. Liu and G. Xu, "A deterministic approach to blind symbol estimation," *IEEE Signal Processing Lett.*, vol. 1, pp. 205–207, Dec. 1994.
[22] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 516–525, Feb. 1995.
[23] A. Oppenheim and R. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
[24] M. Pagano, "On periodic and multiple autoregressions," *Ann. Statist.*, pp. 1310–1317, 1978.
[25] B. Porat and B. Friedlander, "Blind equalization of digital communication channels using higher-order moments," *IEEE Trans. Signal Processing*, vol. 39, pp. 522–526, Feb. 1991.
[26] J. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1989.
[27] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Trans. Inform. Theory*, vol. 36, pp. 312–321, Mar. 1990.

[28] D. T. M. Slock, "Blind fractionally-spaced equalization, perfect-reconstruction filter banks and multichannel linear prediction," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, vol. IV, Adelaide, Australia, 1994, pp. 585–588.
[29] D. T. M. Slock and C. Papadias, "Further results on blind identification and equalization of multiple FIR channels," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, Detroit, MI, May 1995, pp. 1964–1967.
[30] L. Tong, "Blind sequence estimation," *IEEE Trans. Commun.*, vol. 43, pp. 2986–2994, Dec. 1995.
[31] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inform. Theory*, vol. 40, pp. 340–349, Mar. 1994.
[32] L. Tong, G. Xu, B. Hassibi, and T. Kailath, "Blind channel identification based on second-order statistics: A frequency-domain approach," *IEEE Trans. Inform. Theory*, vol. 41, pp. 329–334, Jan. 1995.
[33] M. Tsatsanis and G. B. Giannakis, "Beamforming techniques for multiuser detection in CDMA systems," *Proc. 29th Conf. Inform. Sci. Syst.*, The Johns Hopkins Univ., Baltimore, MD, Mar. 1995, pp. 377–381.
[34] J. Tugnait, "On blind identifiability of multipath channels using fractional sampling and second-order cyclostationary statistics," *IEEE Trans. Inform. Theory*, vol. 41, pp. 308–311, Jan. 1995.
[35] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Trans. Signal Processing*, vol. 43, pp. 2982–2993, Dec. 1995.

Georgios B. Giannakis (F'97) received the Diploma in electrical engineering from the National Technical University of Athens, Athens, Greece, in 1981. From September 1982 to July 1986, he was with the University of Southern California (USC), Los Angeles, where he received the MSc. degree in electrical engineering in 1983, the MSc. degree in mathematics in 1986, and the Ph.D. degree in electrical engineering in 1986.

After lecturing for one year at USC, he joined the University of Virginia, Charlottesville, in September 1987, where he is now a Professor in the Department of Electrical Engineering. His general interests lie in the areas of signal processing, estimation and detection theory, and system identification. Specific research areas of current interest include diversity techniques for channel estimation and multiuser communications, nonstationary and cyclostationary signal analysis, wavelets in statistical signal processing, and non-Gaussian signal processing using higher order statistics with applications to sonar, array, and image processing.

Dr. Giannakis received the IEEE Signal Processing Society's 1992 Paper Award in the Statistical Signal and Array Processing (SSAP) area. He co-organized the 1993 IEEE Signal Processing Workshop on Higher Order Statistics, the 1996 IEEE Workshop on Statistical Signal and Array Processing, and the first IEEE Signal Processing Workshop on Wireless Communications in 1997. He guest (co-)edited two special issues on high-order statistics (*International Journal of Adaptive Control and Signal Processing* and the EURASIP journal *Signal Processing*) and a special issue on signal processing for advanced communications (IEEE TRANSACTIONS ON SIGNAL PROCESSING). He has served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE SIGNAL PROCESSING LETTERS, a secretary of the Signal Processing Conference Board, and a member of the SP Publications board and the SSAP Technical Committee. He is also a member of the IMS and the European Association for Signal Processing.

Steven D. Halford (S'86) received the B.S. and M.S. degrees in electrical engineering from the Georgia Institute of Technology, Atlanta, in 1988 and 1990, respectively. In January 1997, he received the Ph.D. degree in electrical engineering from the University of Virginia, Charlottesville, where he was a Graduate Research Assistant in the Communications Systems Laboratory.

From 1988 until 1990, he was employed by Southern Bell Telephone Company, Atlanta, where he worked on narrowband ISDN. From 1993 to 1995, he was with the Naval Research Laboratories, Washington, DC, working on ATM for wireless communications. He is now with the Government Communications Systems Division, Harris Corporation, Melbourne, FL. His current research interests are in the general areas of channel equalization, signal extraction, and signal processing to improve wireless communications.