Performance Analysis of Godard-Based Blind Channel Identification

Philip Schniter, *Member, IEEE*, Raúl A. Casas, Azzédine Touzni, *Associate Member, IEEE*, and C. Richard Johnson, Jr., *Fellow, IEEE*

***Abstract*—We analyze a blind channel impulse response identification scheme based on the cross correlation of blind symbol estimates with the received signal. The symbol estimates specified are those minimizing the Godard (or constant modulus) criterion, for which mean-squared symbol estimation error bounds have recently been derived. In this paper, we derive upper bounds for the average squared parameter estimation error (ASPE) of the blind identification scheme that depend on the mean-squared error of the Wiener equalizer, the kurtoses of the desired and interfering sources, and the channel impulse response. The effects of finite data length and stochastic gradient equalizer design on ASPE are also investigated. All results are derived in a general multiuser vector-channel context.**

***Index Terms*—Blind channel identification, blind deconvolution, constant modulus algorithm, Godard algorithm.**

I. Introduction

CONSIDER the linear system of Fig. 1, where a desired source sequence $\left\{s_n^{(0)}\right\}$ combines linearly with K interferers through vector channels $\left\{\boldsymbol{h}^{(0)}(z),\ldots,\boldsymbol{h}^{(K)}(z)\right\}$. Our goal is to estimate the impulse response coefficients of the linear channel $\left\{\boldsymbol{h}^{(0)}(z)\right\}$ knowing only the statistics of the received signal $\{\boldsymbol{r}_n\}$. The literature refers to this problem as *blind* channel identification [1].

In this paper, we analyze the performance of the blind channel identification scheme illustrated in Fig. 2, whereby blind symbol estimates $\{y_n\}$ are cross-correlated with the received signal $\{\boldsymbol{r}_n\}$ under the presumption that the source processes $\left\{s_n^{(0)}\right\}\cdots\left\{s_n^{(K)}\right\}$ are mutually independent. We focus specifically on the case of blind linear symbol estimation according to the Godard, or constant modulus (CM), criterion [2]–[4].

There are various applications of the proposed adaptive channel identification algorithm. For one, engineers need a detailed understanding of expected channel characteristics when designing a communication system. The identification algorithm could be used "off-line" during such a design phase to yield channel estimates with guaranteed performance. In addition, the tracking capabilities of the algorithm would allow a study of channel time-variations (e.g., [5]). Perhaps more interesting are "on-line" applications, where the proposed identification algorithm would be used in an operating receiver. One such idea is that accurate channel estimates, generated using a low-complexity adaptive linear equalizer, could be used by a complex and high-performance symbol detector such as a maximum-likelihood sequence detector (MLSD). Alternatively, the blind channel estimates could be used to initialize a blind decision feedback equalizer or other equalization scheme for which cold blind startup is not feasible [6].

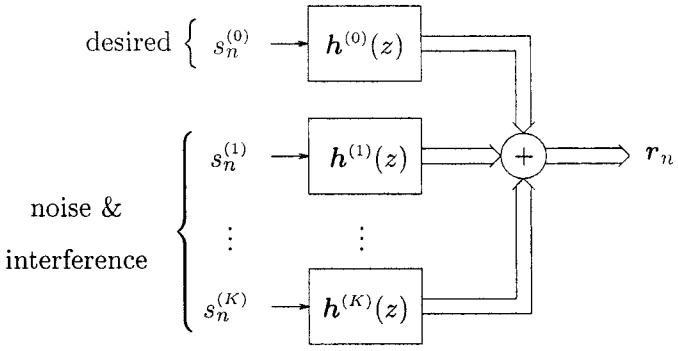

Fig. 1. Linear system model with K sources of interference.

Linear estimation via the CM criterion has become perhaps the most studied and implemented means of blind symbol estimation for data communication over dispersive channels (see, e.g., [4] and the references within). The popularity of CM methods are usually attributed to

i) the existence of a simple adaptive algorithm ("CMA" [2], [3]) for estimation and tracking of the CM-minimizing linear equalizer $\boldsymbol{f}_{c,\nu}(z)$;
ii) the typically excellent mean-squared error (MSE) performance of CM-minimizing equalizers;
iii) the insensitivity to residual carrier phase/frequency offsets in received signal $\boldsymbol{r}_n$.

The second of these two points was first conjectured in the seminal works [2], [3] and recently established in [7] for the arbitrary linear model[1] of Fig. 1.

In this paper, we derive upper bounds for the average squared parameter error (ASPE) of blind channel parameter estimates

Manuscript received September 24, 1999; revised April 25, 2001. P. Schniter and C. R. Johnson, Jr. were supported in part by the National Science Foundation under Grants MIP-9528363 and ECS-9811297, an Intel Foundation fellowship, and Applied Signal Technology, Inc. A. Touzni was supported by INRIA, Rocquencourt, France. The associate editor coordinating the review of this paper and approving it for publication was Dr. Sergio Barbarossa.

P. Schniter is with Department of Electrical Engineering, The Ohio State University, Columbus, OH 43210 USA (e-mail: schniter@ee.eng.ohio-state.edu).

R. A. Casas and A. Touzni are with NxtWave Communications, Langhorne, PA 19047 USA.

C. R. Johnson, Jr. is with the Department of Electrical Engineering, Cornell University, Ithaca, NY 14853 USA.

Publisher Item Identifier S 1053-587X(01)05865-2.

[1]A (nonclosed-form) bounding procedure for the MSE of CM-minimizing equalizers was presented earlier by Zeng *et al.* [9] under a single-user AWGN channel model.

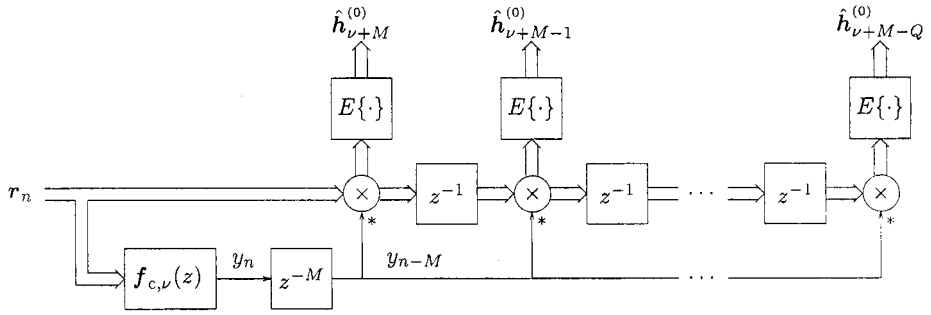

Fig. 2. Blind channel identification using CM-minimizing equalizer $\boldsymbol{f}_{c,\nu}(z)$, which generates nearly MMSE ν-delayed symbol estimates of the zeroth user, i.e., $\{y_n\} \approx \{s^{(0)}_{n-\nu}\}$.

generated by the method of Fig. 2. The bounds are a function of the channel coefficients, the source kurtoses, and the symbol estimation delay. Next, we derive the expected ASPE that results when the correlations in Fig. 2 are estimated from N-length data blocks. Finally, we discuss the effect of stochastic-gradient equalizer design (i.e., the use of CMA versus exact gradient descent) on ASPE. All results are derived in the multiuser vector-channel context of Fig. 1.

The organization of the paper is as follows. Section II discusses the properties of the system model and reviews CM-minimizing symbol estimation and blind channel identification, Section III derives bounds for the MSE performance of the channel estimation scheme, and Section IV covers implementational issues such as finite data effects and the use of CMA. Section V presents the results of numerical simulations verifying our performance analyzes, and Section VI concludes the paper.

II. Background

In this section, we give background information on the linear system model, the CM criterion, and on blind channel identification. The following notation is used throughout:

$(\cdot)^t$ transpose;
$(\cdot)^*$ conjugation;
$(\cdot)^H$ hermitian;
$\mathrm{E}\{\cdot\}$ expectation.

In addition, $(\cdot)^\dagger$ denotes the Moore–Penrose pseudo-inverse, and $||\boldsymbol{x}||_p$ the p-norm defined by $\sqrt[p]{\sum_i |x_i|^p}$. In general, we use boldface lowercase type to denote vector quantities and boldface uppercase type to denote matrix quantities.

A. Linear System Model

First, we formalize the linear time-invariant multi-channel model illustrated in Fig. 1. Say that the desired symbol sequence $\left\{s^{(0)}_n\right\}$ and K sources of interference $\left\{s^{(1)}_n\right\}, \ldots, \left\{s^{(K)}_n\right\}$ each pass through separate linear channels before being observed. The interference processes may correspond, e.g., to co-channel interference signals or additive noise processes.[2] In addition, say that the equalizer uses a sequence of the P-dimensional vector observations $\{\boldsymbol{r}_n\}$ to estimate (a possibly delayed version of) the desired source sequence, where the case $P > 1$ corresponds to the use of multiple sensors and/or sampling at an integer multiple of the symbol rate. The observations can be written

$$\boldsymbol{r}_n = \sum_{k=0}^{K} \sum_{i=0}^{\infty} \boldsymbol{h}^{(k)}_i s^{(k)}_{n-i} \tag{1}$$

where $\left\{\boldsymbol{h}^{(k)}_n\right\}$ denotes the impulse response coefficients of the linear time-invariant (LTI) channel $\boldsymbol{h}^{(k)}(z)$. We assume that $\boldsymbol{h}^{(k)}(z)$ is causal and bounded-input bounded-output (BIBO) stable. Note that such $\boldsymbol{h}^{(k)}(z)$ admit infinite-duration impulse response (IIR) channel models.

As shown in Fig. 2, linear estimates $\{y_n\}$ of $\left\{s^{(k)}_{n-\nu}\right\}$, for fixed estimation delay ν, are generated from the vector-valued observation sequence $\{\boldsymbol{r}_n\}$. Using $\{\boldsymbol{f}_n\}$ to denote the impulse response of the linear equalizer $\boldsymbol{f}(z)$, the symbol estimates are formed as $y_n = \sum_{i=-\infty}^{\infty} \boldsymbol{f}^H_i \boldsymbol{r}_{n-i}$. We will assume that the linear system $\boldsymbol{f}(z)$ is BIBO stable with *constrained* ARMA structure, i.e., certain polynomial coefficients in the numerator and denominator of $\boldsymbol{f}(z)$ may be held at zero.

The global channel-plus-equalizers $q^{(k)}(z) := \boldsymbol{f}^H(z) \boldsymbol{h}^{(k)}(z)$ are often used in the sequel. The impulse response coefficients of $q^{(k)}(z)$ can be written

$$q^{(k)}_n = \sum_{i=-\infty}^{\infty} \boldsymbol{f}^H_i \boldsymbol{h}^{(k)}_{n-i} \tag{2}$$

allowing the estimates to be written as $y_n = \sum_k \sum_i q^{(k)}_i s^{(k)}_{n-i}$. Adopting the following vector notation helps to streamline the remainder of the paper.

$$\begin{aligned}
\underline{\boldsymbol{q}}^{(k)} &:= (\ldots, q^{(k)}_{-1}, q^{(k)}_0, q^{(k)}_1, \ldots)^t \\
\underline{\boldsymbol{q}} &:= \Big(\ldots, q^{(0)}_{-1}, q^{(1)}_{-1}, \ldots, q^{(K)}_{-1}, q^{(0)}_0, q^{(1)}_0, \ldots, q^{(K)}_0, q^{(0)}_1 \\
&\qquad q^{(1)}_1, \ldots, q^{(K)}_1, \ldots\Big)^t \\
\underline{\boldsymbol{s}}^{(k)}(n) &:= \left(\ldots, s^{(k)}_{n+1}, s^{(k)}_n, s^{(k)}_{n-1}, \ldots\right)^t \\
\underline{\boldsymbol{s}}(n) &:= \Big(\ldots, s^{(0)}_{n-1}, s^{(1)}_{n+1}, \ldots, s^{(K)}_{n+1}, s^{(0)}_n, s^{(1)}_n, \ldots, s^{(K)}_n \\
&\qquad s^{(0)}_{n-1}, s^{(1)}_{n-1}, \ldots, s^{(K)}_{n-1}, \ldots\Big)^t.
\end{aligned} \tag{3}$$

[2]Modeling AWGN of variance σ^2_w at P sensors requires P noise sources $\{s^{(k)}_n\}$ with corresponding channels of the form $\boldsymbol{h}^{(k)}(z) = (0, \ldots, 0, \sigma_w/\sigma_s, 0, \ldots, 0)^t$.

For instance, the estimates can be rewritten concisely as

$$y_n = \sum_{k=0}^{K} \underline{\boldsymbol{q}}^{(k)^t} \underline{\boldsymbol{s}}^{(k)}(n) = \underline{\boldsymbol{q}}^t \underline{\boldsymbol{s}}(n). \tag{4}$$

We now point out two important properties of $\underline{\boldsymbol{q}}$. First, recognize that a particular channel and set of equalizer constraints will restrict the set of attainable global responses, which we will denote by $\mathcal{Q}_{\mathrm{a}}$. For example, when the equalizer is finite impulse response (FIR) but otherwise unconstrained, (2) implies that $\underline{\boldsymbol{q}} \in \mathcal{Q}_{\mathrm{a}} = \mathrm{row}(\mathcal{H})$, where

$$\mathcal{H} := \begin{pmatrix} \boldsymbol{h}_0^{(0)} \ldots \boldsymbol{h}_0^{(K)} & \boldsymbol{h}_1^{(0)} \ldots \boldsymbol{h}_1^{(K)} & \boldsymbol{h}_2^{(0)} \ldots \boldsymbol{h}_2^{(K)} & \ldots \\ 0 \ldots 0 & \boldsymbol{h}_0^{(0)} \ldots \boldsymbol{h}_0^{(K)} & \boldsymbol{h}_1^{(0)} \ldots \boldsymbol{h}_1^{(K)} & \ldots \\ \vdots \quad \vdots & \vdots \quad \vdots & \vdots \quad \vdots & \\ 0 \ldots 0 & 0 \ldots 0 & \boldsymbol{h}_0^{(0)} \ldots \boldsymbol{h}_0^{(K)} & \ldots \end{pmatrix}. \tag{5}$$

Restricting the equalizer to be sparse or autoregressive, for example, would generate different attainable sets $\mathcal{Q}_{\mathrm{a}}$. Second, BIBO stable $\boldsymbol{f}(z)$ and $\boldsymbol{h}^{(k)}(z)$ imply BIBO stable $q^{(k)}(z)$ so that $\|\underline{\boldsymbol{q}}^{(k)}\|_p$ exists for all $p \geq 1$, and thus, $\|\underline{\boldsymbol{q}}\|_p$ does as well.

Throughout the paper, we make the following assumptions on the $K+1$ source processes.

- S1) For all k, $\left\{s_n^{(k)}\right\}$ is zero-mean i.i.d.
- S2) $\left\{s_n^{(0)}\right\}, \ldots, \left\{s_n^{(K)}\right\}$ are jointly statistically independent.
- S3) For all k, $\mathrm{E}\left\{|s_n^{(k)}|^2\right\} = \sigma_s^2$.
- S4) $\mathcal{K}\left(s_n^{(0)}\right) < 0$, where $\mathcal{K}(\cdot)$ denotes kurtosis

$$\mathcal{K}(s_n) := \mathrm{E}\left\{|s_n|^4\right\} - 2\left(E\left\{|s_n|^2\right\}\right)^2 - \left|\mathrm{E}\left\{s_n^2\right\}\right|^2. \tag{6}$$

- S5) If, for any k, $q^{(k)}(z)$ or $\left\{s_n^{(k)}\right\}$ is not real-valued, then $\mathrm{E}\left\{s_n^{(k)^2}\right\} = 0$ for all k.

B. Constant Modulus Criterion

The CM (or Godard) criterion, which was introduced independently in [2] and [3], specifies the minimization of the cost functional J_c defined as

$$J_c(y_n) := \mathrm{E}\left\{\left(|y_n|^2 - \gamma\right)^2\right\}. \tag{7}$$

In (7), γ is a positive design parameter known as the "dispersion constant." We are interested in the MSE performance of the CM-minimizing equalizer because it can be directly related to the ASPE of the blind channel estimation scheme in Fig. 2.

Since both symbol power and channel gain are unknown in the "blind" scenario, blind estimators suffer from a gain ambiguity. To ensure that estimator performance evaluation is meaningful in the face of such ambiguity, we base our evaluation on normalized versions of the blind estimators and normalize by the receiver gain $q_\nu^{(0)}$. Given that the estimate y_n can be decomposed into signal and interference terms as $y_n = q_\nu^{(0)} s_{n-\nu}^{(0)} + \bar{\boldsymbol{q}}^t \bar{\boldsymbol{s}}(n)$, where $\bar{\boldsymbol{q}}$ denotes $\underline{\boldsymbol{q}}$ with the $q_\nu^{(0)}$ term removed and $\bar{\boldsymbol{s}}(n)$ denotes $\underline{\boldsymbol{s}}(n)$ with the $s_{n-\nu}^{(0)}$ term removed, the normalized estimate $y_n/q_\nu^{(0)}$ can be referred to as "conditionally unbiased" since $\mathrm{E}\left\{y_n/q_\nu^{(0)} | s_{n-\nu}^{(0)}\right\} = s_{n-\nu}^{(0)}$. The conditionally unbiased MSE (UMSE) associated with y_n, which is an estimate of $s_{n-\nu}^{(0)}$, is then defined as

$$J_{u,\nu}(y_n) := E\left\{\left|\frac{y_n}{q_\nu^{(0)}} - s_{n-\nu}^{(0)}\right|^2\right\}. \tag{8}$$

Substituting the estimate decomposition into (8), we can write the UMSE in terms of the system response $\underline{\boldsymbol{q}}$:

$$J_{u,\nu}(\underline{\boldsymbol{q}}) = \frac{E\left\{|\bar{\boldsymbol{q}}^t \bar{\boldsymbol{s}}(n)|^2\right\}}{\left|q_\nu^{(0)}\right|^2} = \frac{\|\bar{\boldsymbol{q}}\|_2^2}{\left|\boldsymbol{q}_\nu^{(0)}\right|^2}\sigma_s^2 \tag{9}$$

where the second equality invokes assumptions S1)–S3).

For the linear channel model of Fig. 1, it is possible to upper bound the UMSE of CM-minimizing equalizers of delay ν directly in terms of the UMSE of Wiener symbol equalizers of the same delay, i.e., $J_{u,\nu}\left(\underline{\boldsymbol{q}}_{m,\nu}\right)$. Henceforth, we use $\underline{\boldsymbol{q}}_{m,\nu}$ to denote the MMSE global response associated with symbol delay ν. In the FIR case, S1)–S3) imply that $\underline{\boldsymbol{q}}_{m,\nu} = \mathcal{H}^t(\mathcal{H}^*\mathcal{H}^t)^\dagger \mathcal{H}^* \boldsymbol{e}_\nu^{(0)}$, where $\boldsymbol{e}_\nu^{(k)}$ is a column vector with a single nonzero element of value 1 located such that $\underline{\boldsymbol{q}}^t \boldsymbol{e}_\nu^{(k)} = q_\nu^{(k)}$ [8]. A similar expression exists for the IIR case. Before the statement of the bounds, we introduce some additional terminology.

First, we define *normalized* kurtosis κ [not to be confused with $\mathcal{K}(\cdot)$ in (6)]:

$$\kappa_s^{(k)} := \frac{E\left\{\left|s_n^{(k)}\right|^4\right\}}{\left(E\left\{\left|s_n^{(k)}\right|^2\right\}\right)^2}. \tag{10}$$

Under the following definition of κ_g, our results will hold for both real-valued and complex-valued models.

$$\kappa_g := \begin{cases} 3, & \forall\, k,n : \boldsymbol{h}_n^{(k)} \in \mathbb{R}^P \text{ and } s_n^{(k)} \in \mathbb{R} \\ 2, & \text{otherwise.} \end{cases} \tag{11}$$

Note that under S1) and S5), κ_g represents the normalized kurtosis of a Gaussian source. It can be shown that the normalized and un-normalized kurtoses are related by $\mathcal{K}\left(s_n^{(k)}\right) = \left(\kappa_s^{(k)} - \kappa_g\right)\sigma_s^4$ under S3) and S5). Finally, we define the minimum and maximum (normalized) interference kurtoses and two other quantities that appear later.

$$\kappa_s^{\min} := \min_{0 \leq k \leq K} \kappa_s^{(k)}$$

$$\kappa_s^{\max} := \max_{0 \leq k \leq K} \kappa_s^{(k)}$$

$$\rho_{\min} := \frac{\kappa_g - \kappa_s^{\min}}{\kappa_g - \kappa_s^{(0)}}$$

$$\rho_{\max} := \frac{\kappa_g - \kappa_s^{\max}}{\kappa_g - \kappa_s^{(0)}}.$$

Theorem 1: (See [7] for proof.) If Wiener UMSE $J_{u,\nu}\left(\underline{\boldsymbol{q}}_{m,\nu}\right) < J_{\mathrm{o}}\sigma_s^2$, where we have

$$J_0 := \begin{cases} 2\sqrt{(1+\rho_{min})^{-1}} - 1, & \kappa_s^{\max} \le \kappa_g \\ \frac{1-\sqrt{1-\frac{(3-\rho_{\min})(1+\rho_{\max})}{4}}}{\rho_{\max}+\sqrt{1-\frac{(3-\rho_{\min})(1+\rho_{\max})}{4}}}, & \kappa_s^{\max} > \kappa_g, \rho_{\max} \ne -1 \\ \frac{3-\rho_{\min}}{5+\rho_{\min}}, & \kappa_s^{\max} > \kappa_g, \rho_{\max} = -1 \end{cases} \tag{12}$$

the UMSE of CM-minimizing equalizers associated with the same user/delay can be upper bounded as follows:

$$J_{u\nu}\left(\underline{\boldsymbol{q}}_{m,\nu}\right) \le J_{u,\nu}\left(\underline{\boldsymbol{q}}_{c,\nu}\right) \le J_{u,\nu,|_{c,\nu}}^{\max}$$

where we have (13), shown at the bottom of the page. Furthermore, (12) guarantees the existence of a CM-minimizing equalizer associated with this user/delay when $\underline{\boldsymbol{q}}$ is FIR.

It should be noted that Theorem 1 implicitly incorporates the channel and/or equalizer constraints that define $\mathcal{Q}_{\mathrm{a}}$ through its use of the MMSE response $\underline{\boldsymbol{q}}_{m,\nu}(\in \mathcal{Q}_{\mathrm{a}})$. For example, if $\underline{\boldsymbol{q}}_{m,\nu}$ is a MMSE global response constrained to the set of causal IIR equalizers, then the UMSE bound pertains to CM-minimizing global responses $\underline{\boldsymbol{q}}_{c,\nu}$ obeying the same causal-IIR constraint. As another example, if $\underline{\boldsymbol{q}}_{m,\nu}$ was generated with the constraint that $\boldsymbol{f}_{m,\nu}(z)$ was FIR of length-N_f, then (13) would bound the CM-UMSE for length-N_f delay-ν CM-minimizing equalizers $\boldsymbol{f}_{c,\nu}(z)$.

In typical scenarios a) sub-Gaussian desired source in the presence of AWGN or b) constant-modulus desired source in the presence of nonsuper-Gaussian interference, it turns out that $\rho_{\min} = 1$, simplifying (12) and (13) as [7].

$$\begin{aligned} J_0|_{\rho_{\min}=1} &= \sqrt{2} - 1 \\ J_{u,\nu}|_{c,\nu}^{\max}|_{\rho_{\max}=1} &= \frac{1-\sqrt{2\left(1+\frac{J_{u,\nu}\left(\underline{\boldsymbol{q}}_{m,\nu}\right)}{\sigma_s^2}\right)^{-2}-1}}{1+\sqrt{2\left(1+\frac{J_{u,\nu}\left(\underline{\boldsymbol{q}}_{m,\nu}\right)}{\sigma_s^2}\right)^{-2}-1}}\sigma_s^2 \\ &= J_{u,\nu}\left(\underline{\boldsymbol{q}}_{m,\nu}\right) + \frac{1}{2\sigma_s^2}J_{u,\nu}^2\left(\underline{\boldsymbol{q}}_{m,\nu}\right) \\ &\quad + \mathcal{O}\left(J_{u,\nu}^3\left(\underline{\boldsymbol{q}}_{m,\nu}\right)\right). \end{aligned}$$

C. Channel Identification

Fig. 2 illustrates the proposed blind channel impulse response identification scheme, whereby M-delayed versions of the CM-minimizing symbol estimates $\{y_n\} \approx \left\{s_{n-\nu}^{(0)}\right\}$ are cross-correlated with the vector received samples $\{\boldsymbol{r}_{n-Q},\ldots,\boldsymbol{r}_n\}$, yielding the vector channel parameter estimates $\left\{\hat{\boldsymbol{h}}_{\nu+M-Q}^{(0)},\ldots,\hat{\boldsymbol{h}}_{\nu+M}^{(0)}\right\}$. The δth parameter estimate $\hat{\boldsymbol{h}}_{\nu+M-\delta}^{(0)}$ can be expressed as a scaled version of the true parameter corrupted by an error term

$$\begin{aligned} &\hat{h}_{\nu+M-\delta}^{(0)} \\ &\quad = \mathrm{E}\left\{\boldsymbol{r}_{n-\delta}y_{n-M}^*\right\} \\ &\quad = \mathrm{E}\left\{\sum_{k,j}\boldsymbol{h}_j^{(k)}s_{n-\delta-j}^{(k)}\sum_{l,i}q_i^{(l)*}s_{n-M-i}^{(l)*}\right\} \\ &\quad = \sigma_s^2\sum_{k,i}\boldsymbol{h}_{i+M-\delta}^{(k)}q_i^{(k)*} \\ &\quad = \underbrace{\sigma_s^2 q_\nu^{(0)*}}_{\text{scaling}}\left(\boldsymbol{h}_{\nu+M-\delta}^{(0)} + \underbrace{\frac{\sum_{(k,i)\ne(0,\nu)}\boldsymbol{h}_{i+M-\delta}^{(k)}q_i^{(k)*}}{q_\nu^{(0)*}}}_{\text{error}}\right). \end{aligned} \tag{14}$$

We note that the identification scheme in Fig. 2 bears similarity to the Gooch–Harp method of channel identification [10] illustrated in Fig. 3, whereby the CM-minimizing estimates $\{y_n\}$ are processed by a hard decision device $\mathcal{D}$ before cross correlation. Due to the nonlinear operation $\mathcal{D}$, however, performance analysis of the Gooch–Harp scheme is difficult unless perfect decision-making (i.e., $d_n = s_{n-\nu}$) is assumed. In addition, forming reliable decisions requires carrier phase synchronization (an issue with passband data transmission [11]), which is not required in the identification scheme of Fig. 2.

Many other methods of blind channel identification have been proposed [1], most of which estimate channel coefficients from the observed data directly, i.e., without first forming blind symbol estimates. When there are multiple subchannels $(P > 1)$ satisfying certain conditions, it is possible to accomplish blind identification using only the second-order statistics

$$J_{u,\nu}|_{c,\nu}^{\max} := \begin{cases} \frac{1-\sqrt{(1+\rho_{\min})\left(1-\frac{J_{u,\nu}\left(\underline{q}_{m,\nu}\right)}{\sigma_s^2}\right)^{-2}-\rho_{\min}}}{\rho_{\min}+\sqrt{(1+\rho_{\min})\left(\frac{1+J_{u,\nu}\left(\underline{q}_{m,\nu}\right)}{\sigma_s^2}\right)^{-2}-\rho_{\min}}}\sigma_s^2, & \kappa_s^{\max} \le \kappa_g \\ \frac{1-\sqrt{(1+\rho_{\min})\left(1+\frac{J_{u,\nu}\left(\underline{q}_{m,\nu}\right)}{\sigma_s^2}\right)^{-2}\left(1+\rho_{\max}\frac{J_{u,\nu}^2\left(\underline{q}_{m,\nu}\right)}{\sigma_s^4}\right)-\rho_{\min}}}{\rho_{\min}+\sqrt{(1+\rho_{\min})\left(1+\frac{J_{u,\nu}\left(\underline{q}_{m,\nu}\right)}{\sigma_s^2}\right)^{-2}\left(1+\rho_{\max}\frac{J_{u,\nu}^2\left(\underline{q}_{m,\nu}\right)}{\sigma_s^4}\right)-\rho_{\min}}}\sigma_s^2, & \kappa_s^{\max} > \kappa_g \end{cases} \tag{13}$$

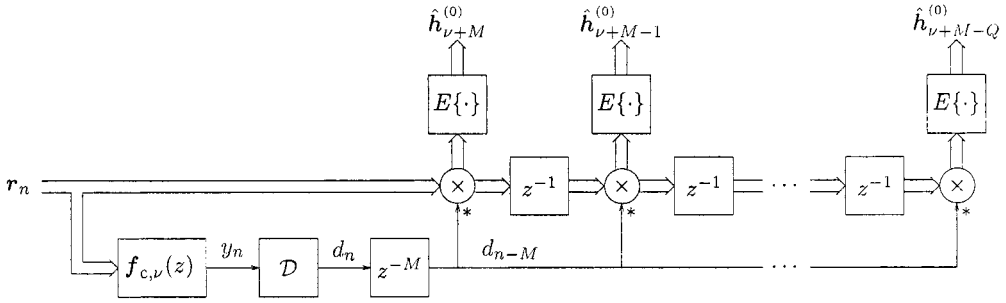

Fig. 3. Gooch–Harp method of blind channel identification.

(SOS) of the observed process (see, e.g., the references in [12]). Most SOS-based techniques, however, are known to fail catastrophically when the channel order is overestimated (see, e.g., the discussion in [13]) or underestimated [14]. An exception is the approach in [13], where, similar to Figs. 2 and 3, the channel coefficients are estimated using cross correlations with blind symbol estimates based on linear prediction. The CM-based schemes in Figs. 2 and 3, however, give good performance, even when the typical channel identifiability conditions fail.

III. Blind Identification—Performance Bounds

We are interested in quantifying the error of the $Q+1$ parameter estimates $\left\{\hat{\boldsymbol{h}}^{(0)}_{\nu+M-Q},\ldots,\hat{\boldsymbol{h}}^{(0)}_{\nu+M}\right\}$ relative to the true parameter subset $\left\{\boldsymbol{h}^{(0)}_{\nu+M-Q},\ldots,\boldsymbol{h}^{(0)}_{\nu+M}\right\}$. We tolerate arbitrary scaling of the total estimated channel response and define our average squared parameter error (ASPE) criterion as follows:

$$\mathcal{E}_{\underline{\hat{h}}} := \min_{\theta\in\mathbb{C}} \frac{1}{Q+1}\sum_{\delta=0}^{Q}\left\|\theta\hat{\boldsymbol{h}}^{(0)}_{\nu+M-\delta}-\boldsymbol{h}^{(0)}_{\nu+M-\delta}\right\|_2^2 \tag{15}$$

$$= \min_{\theta\in\mathbb{C}} \frac{1}{Q+1}\left\|\theta\underbrace{\begin{pmatrix}\hat{\boldsymbol{h}}^{(0)}_{\nu+M}\\ \vdots \\ \hat{\boldsymbol{h}}^{(0)}_{\nu+M-Q}\end{pmatrix}}_{\underline{\hat{\boldsymbol{h}}}^{(0)}_{\nu+M}} - \underbrace{\begin{pmatrix}\boldsymbol{h}^{(0)}_{\nu+M}\\ \vdots \\ \boldsymbol{h}^{(0)}_{\nu+M-Q}\end{pmatrix}}_{\underline{\boldsymbol{h}}^{(0)}_{\nu+M}}\right\|_2^2. \tag{16}$$

Note that by choosing M and Q large enough, an arbitrarily large subset of the total channel response $\left\{\boldsymbol{h}^{(0)}_n\right\}$ may be estimated, regardless of the symbol estimation delay ν.

Theorem 2: For symbol estimation delay ν, the ASPE generated by the blind channel identification scheme in Fig. 2 can be upper bounded as

$$\mathcal{E}_{\underline{\hat{h}}} \le \frac{\|\boldsymbol{\mathcal{H}}_M\|^2 J_{u,\nu}|^{\max}_{c,\nu}}{(Q+1)\sigma_s^2} \tag{17}$$

$$\le \frac{\mathrm{E}\left\{\|\boldsymbol{r}_n\|_2^2\right\} J_{u,\nu}|^{\max}_{c,v}}{\sigma_s^4} \tag{18}$$

when the Wiener equalizer satisfies the UMSE condition $J_{u,\nu}(\underline{\boldsymbol{q}}_{m,\nu})\le J_o\sigma_s^2$ in Theorem 1 . The operator $\boldsymbol{\mathcal{H}}_M$ is defined as

$$\boldsymbol{\mathcal{H}}_M : l_1\left(\mathbb{C}^P\right)\to\mathbb{C}^{PQ}\ \text{s.t.}\ \boldsymbol{\mathcal{H}}_M\underline{\boldsymbol{q}} = \begin{pmatrix}\sum_{k,i}\boldsymbol{h}^{(k)}_{i+M}q^{(k)}_i\\ \sum_{k,i}\boldsymbol{h}^{(k)}_{i+M-1}q^{(k)}_i\\ \vdots\\ \sum_{k,i}\boldsymbol{h}^{(k)}_{i+M-Q}q^{(k)}_i\end{pmatrix} \tag{19}$$

with induced norm

$$\|\boldsymbol{\mathcal{H}}_M\| := \sup_{\underline{q}\neq\underline{0}}\frac{\left\|\boldsymbol{\mathcal{H}}_M\underline{\boldsymbol{q}}^*\right\|_2}{\left\|\underline{\boldsymbol{q}}\right\|_2}. \tag{20}$$

Recall that $\underline{\boldsymbol{q}}$ and $l_1\left(\mathbb{C}^P\right)$ were defined in Section II-A. The operator $\boldsymbol{\mathcal{H}}_M$ is a generalization of an M-shifted version of $\boldsymbol{\mathcal{H}}$ from (5) and is employed so that our bound applies to both finite and infinite-dimensional channels and equalizers. When $\boldsymbol{h}^{(k)}(z)$ are FIR, $\boldsymbol{\mathcal{H}}_M$ reduces to a block Toeplitz matrix, and $\|\boldsymbol{\mathcal{H}}_M\|$ specifies its largest singular value.

Proof: First, we define

$$\bar{\boldsymbol{\mathcal{H}}}_M : l_1\left(\mathbb{C}^P\right)\to\mathbb{C}^{PQ}\quad\text{s.t.}\ \bar{\boldsymbol{\mathcal{H}}}_M\bar{\boldsymbol{q}} = \begin{pmatrix}\sum_{(k,i)\neq(0,\nu)}\boldsymbol{h}^{(k)}_{i-M}q^{(k)}_i\\ \sum_{(k,i)\neq(0,\nu)}\boldsymbol{h}^{(k)}_{i-M-}q^{(k)}_i\\ \vdots\\ \sum_{(k,i)\neq(0,\nu)}\boldsymbol{h}^{(k)}_{i+M-Q}q^{(k)}_i\end{pmatrix}. \tag{21}$$

$\bar{\boldsymbol{\mathcal{H}}}_M$ is a version of $\boldsymbol{\mathcal{H}}_M$ with the components for the zeroth source at delay ν removed, and $\bar{\boldsymbol{q}}$ is a version of $\underline{\boldsymbol{q}}$ with the $q^{(0)}_\nu$ element extracted.

Using (14) and the definitions of $\underline{\hat{\boldsymbol{h}}}^{(0)}_{\nu+M}$ and $\underline{\boldsymbol{h}}^{(0)}_{\nu+M}$ in (16), the operators $\boldsymbol{\mathcal{H}}_M$ and $\bar{\boldsymbol{\mathcal{H}}}_M$ allow us to write

$$\begin{aligned}\underline{\hat{\boldsymbol{h}}}^{(0)}_{\nu+M} &= \boldsymbol{\mathcal{H}}_M\underline{\boldsymbol{q}}^*\sigma_s^2\\ &= \bar{\mathcal{H}}_M\bar{\boldsymbol{q}}^*\sigma_s^2 + \underline{\boldsymbol{h}}^{(0)}_{\nu+M}q^{(0)*}_\nu\sigma_s^2.\end{aligned} \tag{22}$$

Choosing $\theta = (q^{(0)*}_\nu\sigma_s^2)^{-1}$ in (16)

$$\mathcal{E}_{\underline{\hat{h}}} \le \frac{1}{Q+1}\left\|\frac{\underline{\hat{\boldsymbol{h}}}^{(0)}_{\nu+M}}{q^{(0)*}_\nu\sigma_s^2} - \underline{\boldsymbol{h}}^{(0)}_{\nu+M}\right\|_2^2 = \frac{1}{Q+1}\frac{\|\bar{\mathcal{H}}_M\bar{\boldsymbol{q}}^*\|_2^2}{\left|q^{(0)}_\nu\right|^2}. \tag{23}$$

The induced norm (20) allows further bounding of (23)

$$\mathcal{E}_{\underline{\hat{h}}} \le \frac{1}{Q+1}\frac{\|\bar{\mathcal{H}}_M\|^2\|\bar{\boldsymbol{q}}\|_2^2}{\left|q^{(0)}_\nu\right|^2}. \tag{24}$$

Recalling the definition of UMSE in (9), (24) becomes

$$\mathcal{E}_{\underline{\hat{h}}} \le J_{u,\nu}\left(\underline{\boldsymbol{q}}_{c,\nu}\right)\frac{\|\bar{\mathcal{H}}_M\|^2}{(Q+1)\sigma_s^2}. \tag{25}$$

Since

$$\begin{aligned}\|\mathcal{H}_M\|^2 &= \sup_{\|\underline{\boldsymbol{q}}\|_2=1} \underline{\boldsymbol{q}}^t\mathcal{H}_M^H\mathcal{H}_M\underline{\boldsymbol{q}}^* \ge \sup_{\|\underline{\boldsymbol{q}}\|_2=1, q_\nu^{(0)}=0} \underline{\boldsymbol{q}}^t\mathcal{H}_M^H\mathcal{H}_M\underline{\boldsymbol{q}}^* \\ &= \sup_{\|\bar{\boldsymbol{q}}\|_2=1} \bar{\boldsymbol{q}}^t\bar{\mathcal{H}}_M^H\bar{\mathcal{H}}_M\bar{\boldsymbol{q}}^* = \|\mathcal{H}_M\|^2\end{aligned}$$

(25) yields

$$\mathcal{E}_{\underline{\hat{h}}} \le J_{u,\nu}\left(\underline{\boldsymbol{q}}_{c,\nu}\right)\frac{\|\mathcal{H}_M\|^2}{(Q+1)\sigma_s^2}. \tag{26}$$

When $\left\{\boldsymbol{h}^{(k)}(z)\right\}$, $\mathcal{Q}_\mathrm{a}$, and ν are such that $J_{u,\nu}\left(\underline{\boldsymbol{q}}_{m,\nu}\right) \le J_\mathrm{o}\sigma_s^2$ for J_o in (12), Theorem 1 allows upper bounding of $J_{u,\nu}\left(\underline{\boldsymbol{q}}_{c,\nu}\right)$, and (26) becomes (17).

Simplification of (17) is possible using the fact that

$$\begin{aligned}\|\mathcal{H}_M\underline{\boldsymbol{q}}^*\|_2^2 &= \sum_{\delta=0}^{Q}\left\|\sum_{k,i}\boldsymbol{h}_{i+M-\delta}^{(k)}q_i^{(k)*}\right\|_2^2 \\ &\le \sum_{\delta=0}^{Q}\left(\sum_{k,i}\left\|\boldsymbol{h}_{i+M-\delta}^{(k)}q_i^{(k)*}\right\|_2^2\right) \\ &= \sum_{\delta=0}^{Q}\sum_{k,i}\left\|\boldsymbol{h}_{i+M-\delta}^{(k)}\right\|_2^2|q_i^{(k)}|^2 \\ &\le (Q+1)\sum_{k,i}\left\|\boldsymbol{h}_i^{(k)}\right\|_2^2\|\underline{\boldsymbol{q}}\|_2^2\end{aligned}$$

which implies

$$\|\mathcal{H}_M\|^2 \le (Q+1)\sum_{k,i}\left\|\boldsymbol{h}_i^{(k)}\right\|_2^2. \tag{27}$$

Rewriting (27) using

$$\sum_{k,i}\left\|\boldsymbol{h}_i^{(k)}\right\|_2^2 = \frac{1}{\sigma_s^2}\mathrm{E}\left\{\left\|\sum_{k,i}\boldsymbol{h}_i^{(k)}s_{n-i}^{(k)}\right\|_2^2\right\} = \frac{1}{\sigma_s^2}\mathrm{E}\left\{\|\boldsymbol{r}_n\|_2^2\right\}$$

gives

$$\|\mathcal{H}_M\|_2^2 \le \frac{Q+1}{\sigma_s^4}\mathrm{E}\left\{\|\boldsymbol{r}_n\|_2^2\right\} \tag{28}$$

which leads to (18).

Theorem 2 gives an upper bound for the ASPE that is proportional to the norm of the channel operator[3] and the UMSE of the ν-delayed Wiener equalizer [through the definition of $J_{u,\nu}|_{c,\nu}^{\max}$ in (13)], as well as a looser bound that is proportional to the received power and the Wiener UMSE. Section V plots these upper bounds for comparison with the actual ASPE attained using the CM-minimizing equalizer.

IV. Blind Identification Issues in Practical Implementation

A. ASPE with Finite-Data Correlation Approximation

In practice, the expectation operations in Fig. 2 will be replaced by some sort of block or exponential averages. In this section, we analyze the effect of block averaging on the parameter estimation error. The δth block parameter estimate is defined below for block size N.

$$\hat{\hat{\boldsymbol{h}}}_{\nu+M-\delta}^{(0)} := \frac{1}{N}\sum_{n=0}^{N-1}\boldsymbol{r}_{n-\delta}y_{n-M}^*. \tag{29}$$

Lemma 1: The expected ASPE using N-block estimates of the autocorrelations can be written as (30), shown at the bottom of the page, where the last term in (30) appears only in the case of real-valued sources, channels, and equalizers.[4]

[3]The norm of the channel operator equals the maximum singular value of a suitably defined channel matrix when channel and equalizer are FIR.

[4]Recall assumption S5).

$$\begin{aligned}E\left\{\mathcal{E}_{\underline{\hat{\hat{h}}}}\right\} = \mathcal{E}_{\bar{h}} &+ \frac{\sigma_s^4}{N(Q+1)}\frac{\left|\hat{\hat{\boldsymbol{h}}}_{\nu+M}^{(0)H}\underline{\boldsymbol{h}}_{\nu+M}^{(0)}\right|^2}{\left\|\underline{\hat{\boldsymbol{h}}}_{\nu+M}^{(0)}\right\|_2^4} \\ &\times\sum_{\delta=0}^{Q}\left(\sum_k\left\|\sum_i\boldsymbol{h}_{M-\delta+i}^{(k)}q_i^{(k)*}\right\|^2\left(\kappa_s^{(k)}-1\right)\right) \\ &+\frac{1}{N}\sum_{n,m=0}^{N-1}\left(\sum_{k,j}\boldsymbol{h}_{m-n+j}^{(k)H}\boldsymbol{h}_j^{(k)}\right)\left(\sum_{(c,i)\neq(k,j-M+\delta)}q_i^{(c)*}q_{m-n+i}^{(c)}\right) \\ &+\underbrace{\frac{1}{N^2}\sum_{n,m=0}^{N-1}\left(\sum_{l,i}\boldsymbol{h}_{m-n+M-\delta+i}^{(l)H}q_i^{(l)*}\right)\left(\sum_{(k,d)\neq(l,m-n+i)}\boldsymbol{h}_{n-m-M-\delta+d}^{(k)}q_d^{(k)}\right)}_{\text{iff } s\in\mathbb{R}}\end{aligned} \tag{30}$$

Proof: The expected block-ASPE can be written

$$\begin{aligned}
&\mathrm{E}\left\{\mathcal{E}_{\underline{\hat{h}}}\right\} \\
&\quad = \mathrm{E}\left\{\frac{1}{Q+1}\sum_{\delta=0}^{Q}\left\|\theta_{\underline{\hat{h}}}\hat{\underline{\boldsymbol{h}}}^{(0)}_{\nu+M-\delta} - \boldsymbol{h}^{(0)}_{\nu+M-\delta}\right\|_2^2\right\} \\
&\quad = \mathrm{E}\Bigg\{\frac{1}{Q+1}\sum_{\delta=0}^{Q} \\
&\qquad \times\left(\left\|\theta_{\underline{\hat{h}}}\hat{\underline{\boldsymbol{h}}}^{(0)}_{\nu+M-\delta} - \theta_{\underline{\hat{h}}}\mathrm{E}\left\{\hat{\underline{\boldsymbol{h}}}^{(0)}_{\nu+M-\delta}\right\}\right\|_2^2\right. \\
&\qquad \left.+\left\|\theta_{\underline{\hat{h}}}\mathrm{E}\left\{\hat{\underline{\boldsymbol{h}}}^{(0)}_{\nu+M-\delta}\right\} - \boldsymbol{h}^{(0)}_{\nu+M-\delta}\right\|_2^2\right)\Bigg\}
\end{aligned}$$

where $\theta_{\underline{\hat{h}}}$ is the value of θ minimizing (16). Zeroing the partial derivative of (16) with respect to θ, it is straightforward to show that

$$\theta_{\underline{\hat{h}}} = \frac{\hat{\underline{\boldsymbol{h}}}^{(0)H}_{\nu+M}\underline{\boldsymbol{h}}^{(0)}_{\nu+M}}{\left\|\hat{\underline{\boldsymbol{h}}}^{(0)}_{\nu+M}\right\|_2^2}. \tag{31}$$

From (29) it follows that $\mathrm{E}\left\{\hat{\hat{\boldsymbol{h}}}^{(0)}_{\nu+M-\delta}\right\} = \mathrm{E}\{\boldsymbol{r}_{n-\delta}y^*_{n-M}\} = \hat{\boldsymbol{h}}^{(0)}_{\nu+M-\delta}$, and thus

$$\begin{aligned}
\mathrm{E}\left\{\mathcal{E}_{\underline{\hat{h}}}\right\} &= \frac{1}{Q+1} \\
&\quad\times\sum_{\delta=0}^{Q}\left(\left|\theta_{\underline{\hat{h}}}\right|^2\mathrm{E}\left\{\left\|\hat{\hat{\boldsymbol{h}}}^{(0)}_{\nu+M-\delta} - \hat{\boldsymbol{h}}^{(0)}_{\nu+M-\delta}\right\|_2^2\right\}\right. \\
&\qquad\left.+\left\|\theta_{\underline{\hat{h}}}\hat{\boldsymbol{h}}^{(0)}_{\nu+M-\delta} - \boldsymbol{h}^{(0)}_{\nu+M-\delta}\right\|_2^2\right) \\
&= \mathcal{E}_{\underline{\hat{h}}} + \frac{\left|\theta_{\underline{\hat{h}}}\right|^2}{Q+1}\sum_{\delta=0}^{Q}\boldsymbol{E}\left\{\left\|\hat{\hat{\boldsymbol{h}}}^{(0)}_{\nu+M-\delta} - \hat{\boldsymbol{h}}^{(0)}_{\nu+M-\delta}\right\|_2^2\right\}
\end{aligned} \tag{32}$$

Furthermore

$$\begin{aligned}
\mathrm{E}\left\{\left\|\hat{\hat{\boldsymbol{h}}}^{(0)}_i - \hat{\boldsymbol{h}}^{(0)}_i\right\|_2^2\right\} &= \mathrm{E}\Bigg\{\left\|\hat{\hat{\boldsymbol{h}}}^{(0)}_i\right\|_2^2 - \left(\hat{\hat{\boldsymbol{h}}}^{(0)}_i\right)^H\hat{\boldsymbol{h}}^{(0)}_i \\
&\qquad - \left(\hat{\boldsymbol{h}}^{(0)}_i\right)^H\hat{\hat{\boldsymbol{h}}}^{(0)}_i + \left\|\hat{\boldsymbol{h}}^{(0)}_i\right\|_2^2\Bigg\} \\
&= \mathrm{E}\left\{\left\|\hat{\hat{\boldsymbol{h}}}^{(0)}_i\right\|_2^2\right\} - \mathrm{E}\left\{\hat{\hat{\boldsymbol{h}}}^{(0)}_i\right\}^H\hat{\boldsymbol{h}}^{(0)}_i \\
&\qquad - \left(\hat{\boldsymbol{h}}^{(0)}_i\right)^H\mathrm{E}\left\{\hat{\hat{\boldsymbol{h}}}^{(0)}_i\right\} + \left\|\hat{\boldsymbol{h}}^{(0)}_i\right\|_2^2 \\
&= \mathrm{E}\left\{\left\|\hat{\hat{\boldsymbol{h}}}^{(0)}_i\right\|_2^2\right\} - \left\|\hat{\boldsymbol{h}}^{(0)}_i\right\|_2^2.
\end{aligned} \tag{33}$$

Now, we wish to examine

$$\mathrm{E}\left\{\left\|\hat{\hat{\boldsymbol{h}}}^{(0)}_{\nu+M-\delta}\right\|_2^2\right\} = \sum_{p=1}^{P}\mathrm{E}\left\{\left|\left[\hat{\hat{\boldsymbol{h}}}^{(0)}_{\nu+M-\delta}\right]_p\right|^2\right\}$$

where $[\boldsymbol{h}]_p$ denotes the pth component of $\boldsymbol{h}$. To avoid complicated notation, we focus temporarily on the case $P = 1$. Thus

$$\begin{aligned}
&\mathrm{E}\left\{\left|\hat{\hat{h}}^{(0)}_{\nu-M-\delta}\right|^2\right\} \\
&\quad = \mathrm{E}\left\{\left|\frac{1}{N}\sum_{n=0}^{N-1}\sum_{k,j}h^{(k)}_j s^{(k)}_{n-\delta-j}\sum_{l,i}q^{(l)*}_i s^{(l)*}_{n-M-i}\right|^2\right\} \\
&\quad = \frac{1}{N^2}\sum_{k,j,l,i}h^{(k)}_j q^{(l)*}_i\sum_{a,b,c,d}h^{(a)*}_b q^{(c)}_d \\
&\qquad\times\sum_{n=0}^{N-1}\sum_{m=0}^{N-1}\underbrace{E\left\{s^{(k)}_{n-\delta-j}s^{(l)*}_{n-M-i}s^{(a)*}_{m-\delta-b}s^{(c)}_{m-M-d}\right\}}_{B}.
\end{aligned}$$

The quantity B above vanishes unless the indices k, j, l, i, a, b, c and d are suitably aligned. After a bit of algebra,[5] we have the equation at the bottom of the page, where the last term only appears in the real-valued case. Generalizing the previous expression to the case $P > 1$ and combining the result with (31)–(33), we arrive at (30). ∎

Simulations suggest that for CM-minimizing equalizers $\boldsymbol{f}_{c,\nu}(z)$ and typical values of N, the second term in (30) dominates the first. This implies that the performance of the proposed channel estimation scheme is, in practice, limited by the finite-data correlations rather than by the performance of the blind equalizers. The plots in Section V agree with this notion: Improvement in symbol estimates gained through

[5]Contact the corresponding author for the details.

$$\begin{aligned}
\mathrm{E}\left\{\left|\hat{\hat{h}}^{(0)}_{\nu+M-\delta}\right|^2\right\} &= \left|\hat{h}^{(0)}_{\nu+M-\delta}\right|^2 + \frac{1}{N}\sum_k\left|\sum_i h^{(k)}_{M-\delta+i}q^{(k)*}_i\right|^2\left(\kappa^{(k)}_s - 1\right)\sigma_s^4 \\
&\quad + \frac{1}{N^2}\sum_{n,m=0}^{N-1}\sum_{k,j}h^{(k)}_j h^{(k)*}_{m-n+j}\sum_{(c,i)\neq(k,j-M+\delta)}q^{(c)*}_i q^{(c)}_{m-n+i}\sigma_s^4 \\
&\quad + \underbrace{\frac{1}{N^2}\sum_{n,m=0}^{N-1}\sum_{l,i}h^{(l)*}_{m-n+M-\delta+i}q^{(l)*}_i\sum_{(k,d)\neq(l,m-n+i)}h^{(k)}_{n-m+M-\delta+d}q^{(k)}_d\sigma_s^4}_{\text{iff } s\in\mathbb{R}}
\end{aligned}$$

quantization of $\{y_n\}$ gives the Gooch–Harp scheme [10] only a minor advantage in ASPE.

B. Stochastic Gradient Estimation of CM Equalizer

Practical implementations of the identification scheme in Fig. 2 will not have knowledge of the exact CM-minimizing equalizer $\boldsymbol{f}_{c,\nu}(z)$. Typically, $\boldsymbol{f}_{c,\nu}(z)$ will be replaced by an iteratively updated approximation to $\boldsymbol{f}_{c,\nu}(z)$ generated by CMA, which attempts stochastic gradient minimization of CM cost [2], [3]. For finite-length $\boldsymbol{f}(z)$, CMA updates the equalizer parameters $\left\{\boldsymbol{f}_0,\ldots,\boldsymbol{f}_{N_f-1}\right\}$ using the following rule (where n denotes the time step):

$$\begin{aligned}\boldsymbol{f}_i(n+1) &= \boldsymbol{f}_i(n)+\mu\boldsymbol{r}_{n-i}y_n^*\left(\gamma-|y_n|^2\right)\\ &\quad 0\le i\le N_f-1.\end{aligned}\tag{34}$$

Similar update rules can be derived for equalizer structures that employ feedback (resulting in an IIR equalizer) [15]. In (34), μ is a small positive stepsize.

The operation of CMA can be considered as a two-stage process. Starting from an initialization $\boldsymbol{f}_i(z)$, the CMA-updated equalizer $\boldsymbol{f}(z)$ first converges to a neighborhood of the exact CM-minimizing equalizer $\boldsymbol{f}_{c,\nu}(z)$ associated with some combination of source k and symbol estimation delay ν. The particular {source, delay}, as well as the time to convergence, depend on the initialization $\boldsymbol{f}_i(z)$. Although various initialization procedures have been proposed (see, e.g., the references in [4]), none are known to work "perfectly" in all situations. Recent work, however, has shown that if the signal to interference-plus-noise (SINR) ratio of the estimates generated by $\boldsymbol{f}_i(z)$ is above a prescribed threshold, then small stepsize CMA will converge to a neighborhood of the CM-minimum $\boldsymbol{f}_{c,\nu}(z)$ associated with the same {source, delay} as $\boldsymbol{f}_i(z)$. For i.i.d. sub-Gaussian sources in the presence of AWGN, the SINR threshold equals 3.8 dB [16].

Once the CMA-updated equalizer parameters have converged to a neighborhood of the local CM minimum $\boldsymbol{f}_{c,\nu}(z)$, averaging theory predicts that the CMA-updated equalizer trajectory converges almost surely and in mean to $\boldsymbol{f}_{c,\nu}(z)$ [17]. In practical situations, however, the CMA-updated equalizer will "jitter" around this local minimum, where the amount of jitter is proportional to the stepsize μ and to the average size of the error term $\boldsymbol{r}_{n-1}^*y_n^*\left(\gamma-|y_n|^2\right)$ in (34). It is possible to derive expressions for the *excess* MSE due to CMA, i.e., the difference between the expected MSE of CMA-generated equalizers and the MSE of CM-minimizing equalizers. For example, CMAs excess MSE resulting from the use of source symbols drawn from a nonconstant-modulus alphabet is characterized in [18]. The simulations in Section V, however, seem to indicate that for typical block sizes N, the effects of finite-data correlation approximation overwhelm the effects of CMA-induced excess symbol estimation error. For this reason, we do not investigate further the CMA-induced error.

Throughout our discussion of blind channel estimation, we have been assuming that the CMA-derived symbol estimates are reasonably accurate, which would seem to require use of a data record long enough to support the convergence time of CMA. Even with very short data records, however, it may be possible to adapt CMA using a repeatedly concatenated version of the same data record (similar to "looped" LMS [19]). Since source independence assumptions S1) and S2) become less valid as record length decreases, however, it is difficult to make solid claims about the convergence of such data-reusing CMA schemes. Although of practical importance, CMA data-reuse lies outside the scope of this paper.

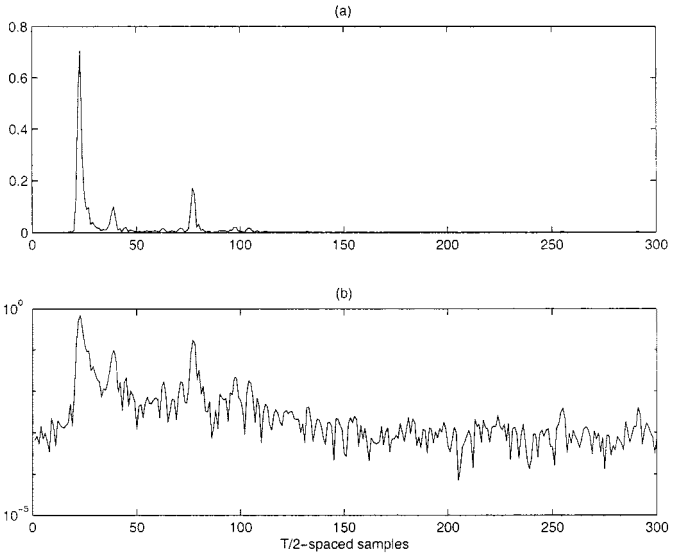

Fig. 4. Magnitude of SPIB channel #3 impulse response shown on (a) linear and (b) log scales.

V. Numerical Examples

Our first experiments were based on complex-valued $T/2$-spaced (i.e., $P=2$) signal processing information base[6] (SPIB) microwave channel response model #3, consisting of 300 $T/2$-spaced samples and depicted in Fig. 4. We blindly estimated the "interesting" subset of the channel response spanning $T/2$-spaced taps 15–85 [see Fig. 8(a) and (b)].

Figs. 5–7 each plot bounds (17) and (18) for the ASPE of the exact CM-minimizing equalizer with exact cross correlations compared with

i) the average ASPE achieved by the proposed CMA-based scheme using block length $N=10^4$;
ii) the average ASPE achieved by the Gooch–Harp scheme [10] using block length N;
iii) the expected ASPE for the exact CM-minimizing equalizer[7] using block length N [from (30)];
iv) the ASPE for the exact CM-minimizing equalizer with exact cross correlations [from (15)];
v) the ASPE for the Wiener equalizer with exact cross correlations [also from (15)].

The following details apply to all experiments: The symbol delay ν was chosen as the MSE-minimizing delay for the

[6]The SPIB microwave channel database resides at http://spib.rice.edu/spib/microwave.html.

[7]The CM-minimizing equalizer $\boldsymbol{f}_{c,\nu}(z)$ was determined numerically using Matlab's "fminunc" routine.

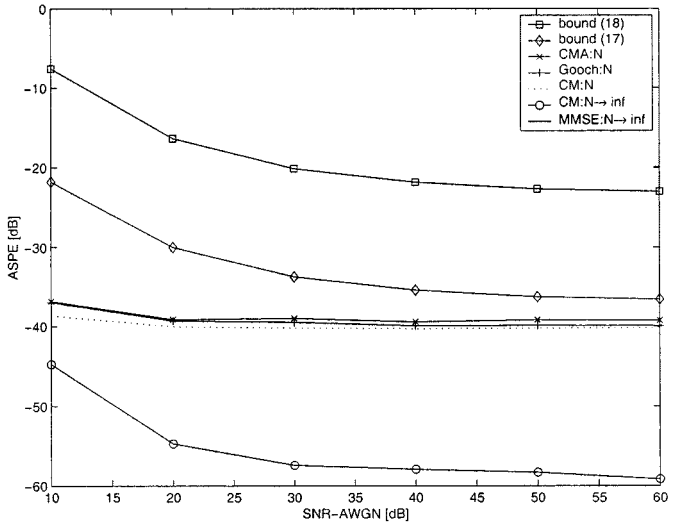

Fig. 5. Average-squared parameter error for $T/2$-spaced SPIB microwave channel #3 versus SNR of AWGN.

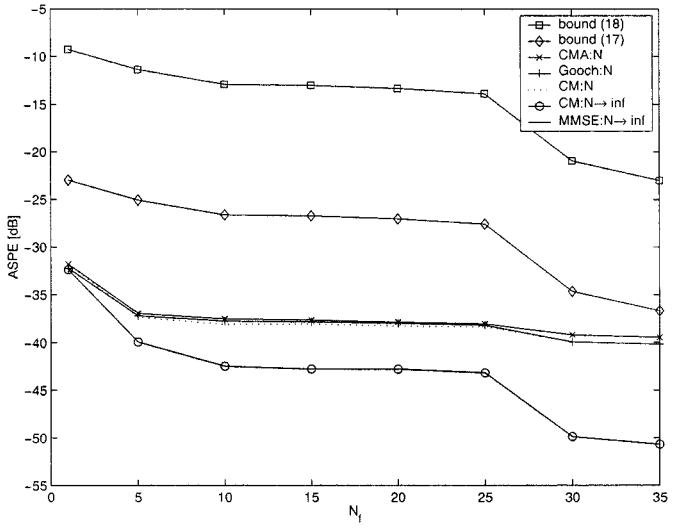

Fig. 6. Average-squared parameter error for $T/2$-spaced SPIB microwave channel #3 versus equalizer length N_f for 64-QAM.

particular combination of channel, noise, and equalizer constraints, and CMA was initialized at $\boldsymbol{f}_{c,\nu}(z)$ and adapted with stepsize $\mu = 10^{-3}$.

In Fig. 5, we use a $T/2$-spaced equalizer with $N_f = 35$ and varied the level of AWGN, and the source alphabet was 64-QAM. The equalizer was long enough to do a good job of "inverting" the channel under high SNR, leading to $\underline{\boldsymbol{g}} \approx \boldsymbol{e}_\nu^{(0)}$ and, thus, accurate channel estimation under perfect cross correlations. Note that the Wiener and CM-minimizing equalizers generated nearly identical performance (i.e., their traces were overlaid), which we expect given that the close relationship implied by Theorem 1. The bounds (17) and (18) lie clearly above the trace corresponding to the exact CM-miminizing equalizer with exact cross correlations, as expected.[8] The dotted line in Fig. 5 corresponds to an exact CM-minimizing

[8]We have determinined that (24) is primarily responsible for the looseness of bound (17), whereas (27) is responsible for the further looseness of (18).

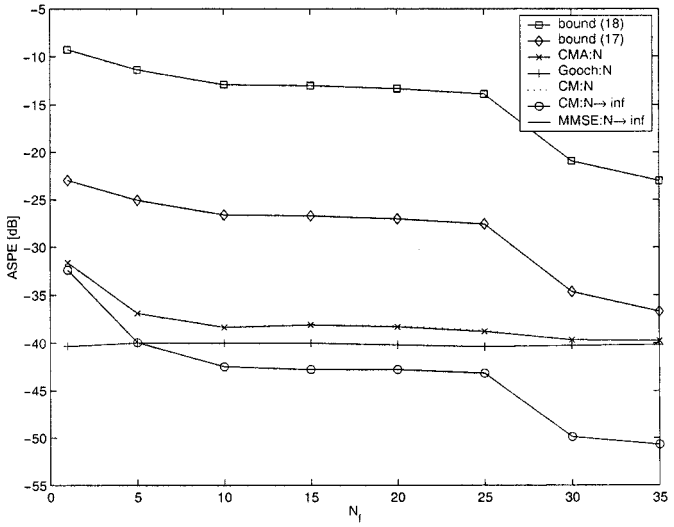

Fig. 7. Average-squared parameter error for $T/2$-spaced SPIB microwave channel #3 versus equalizer length N_f for QPSK.

equalizer with block-averaged cross correlations (for block length $N = 10\,000$); the relative independence from SNR $(> 20$ dB$)$ implies that the N block effects swamp out noise-induced errors. The trace corresponding to the CMA adaptive equalizer and N-block cross correlations is further corrupted by EMSE in the equalizer output (due to a nonconstant modulus 64-QAM source [18] and nonzero stepsize $\mu = 0.001$), although the EMSE contribution to ASPE is minor. In the Gooch–Harp scheme, we process the CMA equalizer output with a nearest-element detector. We expect the equalizer outputs to be cleaned up under high SNR (since decisions should be reliable), although we expect little improvement at low SNR (since decisions should be unreliable given our large 64-QAM alphabet). This behavior is demonstrated in Fig. 5; at low SNR, the decisions are not accurate enough to make Gooch–Harp any better than CMA, whereas at high SNR, the decision device removes residual EMSE in the equalizer output and yields ASPE equivalent to the exact-CM equalizer (which is practically perfect for this equalizer length at high SNR).

In Fig. 6, we examine the effects of insufficient ISI removal by varying the equalizer length from $N_f = 1$ to 40 under a relatively high SNR of 40 dB. All other parameters are those of Fig. 5, and the same trends can be observed. Weiner and CM performance are practically identical, and (17) and (18) loosely bound the ASPE of exact CM equalizers with exact cross correlations. Likewise, the effect of $N = 10\,000$ block size on ASPE swamps out the effect of equalization error for all but the (extremely undermodeled) two-tap equalizer, and performance does not improve much as equalizer length is increased over $N_f = 10$. Finally, CMA-induced EMSE in the equalizer output adds little compared with finite-block-induced error, and the Gooch–Harp scheme effectively removes the EMSE when the equalizer outputs are accurate enough for nearest-element detection but does little good otherwise. Note that the sudden jump in performance for $N_f > 25$ is a result of the equalizer having a long enough time span to compensate for the significant channel feature near tap 75 in Fig. 4.

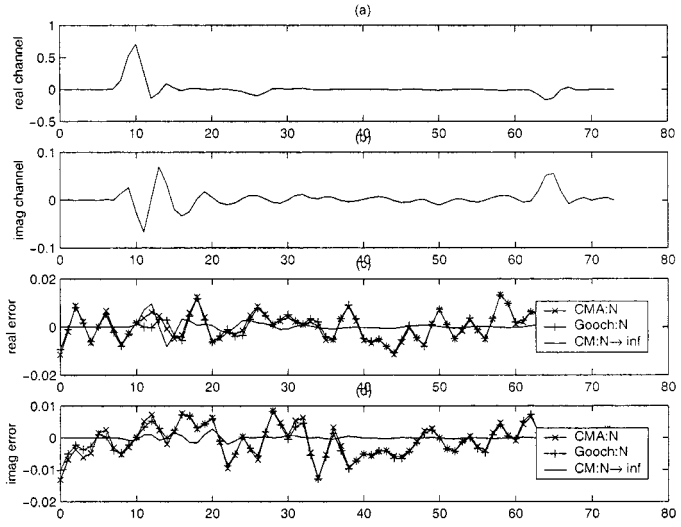

Fig. 8. (a) Real and (b) imaginary components of SPIB microwave channel #3 impulse response spanning taps 15–85. (c) Real and (d) imaginary components of estimation error for SNR $= 40$ dB, $N_f = 35$, and 64-QAM.

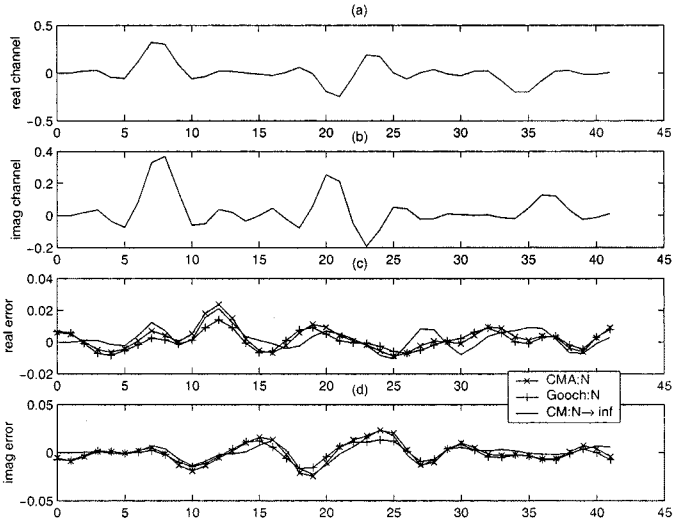

Fig. 9. (a) Real and (b) imaginary components of artificial $T/2$-spaced impulse response. (c) Real and (d) imaginary components of estimation error for SNR $= 40$ dB, $N_f = 28$, and 16-QAM.

Fig. 7 repeats the experiment of Fig. 6 but with a QPSK source. Although the source kurtosis of QPSK is lower than that of 64-QAM, the quantity $\rho_{\min}$, which predicts the UMSE performance of the CM-minimizing equalizer via Theorem 1, does not change. Hence, the ASPE with exact-CM equalization remains identical, as do the bounds (17) and (18). The switch to QPSK does change the performance of CMA and Gooch–Harp adaptive schemes, however. First, the EMSE of CMA is significantly reduced because the source alphabet now has constant modulus [18]. As a result, CMA performance is practically identical to exact-CM performance (assuming N-block cross correlations). The biggest improvement occurs with the Gooch–Harp scheme, however, because (with this channel) reliable QPSK decisions can be made for all equalizer lengths. Hence, N-block cross correlation is the only contributor to Gooch–Harp ASPE in Fig. 7.

Fig. 8 shows typical estimation errors compared with the channel segment being estimated for N-block CMA, N-block Gooch–Harp, and exact-CM equalization with exact cross correlation. In this figure, a 64-QAM source is equalized in 40 dB SNR with an equalizer of $N_f = 40$.

Our second set of experiments were based on the artificially generated complex-valued $T/2$-spaced channel impulse response depicted in Fig. 9(a) and (b) for which we estimated all 42 impulse response coefficients. This channel was generated by filtering a random collection of discrete multipath components with a raised-cosine pulse of rolloff factor 0.2 [11]. The parameters N, μ, and ν are the same as before. Fig. 9(c) and (d) shows typical estimation errors for N-block CMA, N-block Gooch–Harp, and exact-CM equalization with exact cross correlation when $N_f = 28$, SNR $= 40$ dB, and the source is 16-QAM. InFig. 10, we keep these parameter choices but vary the equalizer length from $N_f = 8$ to 28 and plot the resulting ASPE. Fig. 10 demonstrates behavior similar to Figs. 6 and 7. In Fig. 10, however, the effects of finite-block correlation approximation do not dominate the ASPE because the ASPE is well above the "error floor" visible in Figs. 6 and 7. We also

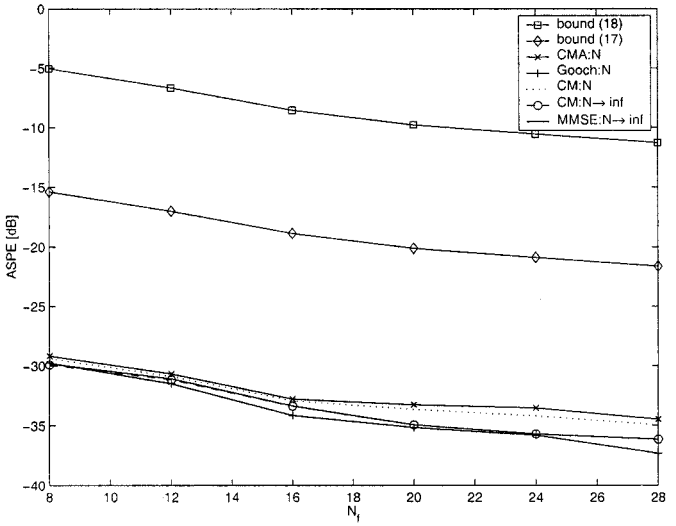

Fig. 10. Average-squared parameter error for artificial $T/2$-spaced channel versus equalizer length N_f for 16-QAM.

note that, in Fig. 10, the relative performance of Gooch–Harp is somewhere between that seen for 64-QAM (in Fig. 6) and QPSK (in Fig. 7) since the source is now 16-QAM.

VI. Conclusions

We have analyzed the performance of a blind channel identification scheme based on the cross correlation of CM-minimizing blind symbol estimates with the received signal. Leveraging recent results on the unbiased MSE of CM-minimizing equalizers, upper bounds on the average squared channel parameter estimation error (ASPE) were derived. Implementational aspects were also considered, such as ASPE increase due to finite-data correlation approximation and stochastic gradient approximations of the CM-minimizing equalizer. Finally, experiments using SPIB microwave channel models were presented to verify the results of our analyses. One conclusion was that the ASPE due to finite-data correlation approximation may overwhelm the ASPE

due to errors in the CMA equalizer output, even for reasonably large data record lengths.

REFERENCES

[1] L. Tong and S. Perreau, "Blind channel estimation: From subspace to maximum likelihood methods," *Proc. IEEE, Special Issue on Blind System Identification and Estimation*, vol. 86, pp. 1951–68, Oct. 1998.

[2] 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.

[3] J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 459–72, Apr. 1983.

[4] C. R. Johnson Jr., P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, and R. A. Casas, "Blind equalization using the constant modulus criterion: A review," *Proc. IEEE, Special Issue on Blind System Identification and Estimation*, vol. 86, pp. 1927–1950, Oct. 1998.

[5] J. D. Behm, T. J. Endres, P. Schniter, M. L. Alberi, C. Prettie, C. R. Johnson, Jr., and I. Fijalkow, "Characterization of an empirically-derived database of time-varying microwave channel responses," in *Proc Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 1997, pp. 1549–1553.

[6] R. A. Casas, T. J. Endres, A. Touzni, C. R. Johnson, Jr., and J. R. Treichler, "Current approaches to blind decision-feedback equalization," in *Signal Processing Advances in Communications*, G. B. Giannakis, P. Stoica, Y. Hua, and L. Tong, Eds. New York: Wiley, 2000, vol. 1.

[7] P. Schniter and C. R. Johnson, Jr., "Bounds for the MSE performance of constant modulus estimators," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2544–2560, Nov. 2000.

[8] P. Schniter, "Blind estimation without priors: Performance, convergence, and efficient implementation," Ph.D. dissertation, Cornell Univ., Ithaca, NY, 2000.

[9] H. H. Zeng, L. Tong, and C. R. Johnson Jr., "An analysis of constant modulus receivers," *IEEE Trans. Signal Processing*, vol. 47, pp. 2990–2999, Nov. 1999.

[10] R. P. Gooch and J. C. Harp, "Blind channel identification using the constant modulus adaptive algorithm," in *Proc IEEE Int. Conf. Commun.*, Philadelphia, PA, June 1988, pp. 75–79.

[11] J. G. Proakis, *Digital Communications*, 2nd ed. New York: McGraw-Hill, 1989.

[12] H. Liu, G. Xu, L. Tong, and T. Kailath, "Recent developments in blind channel equalization: From cyclostationarity to subspaces," *Signal Process.*, vol. 50, pp. 83–99, 1996.

[13] K. Abed-Meraim, E. Moulines, and P. Loubaton, "Prediction Error Method for Second-Order Blind Identification," *IEEE Trans. Signal Processing*, vol. 45, pp. 694–705, Mar. 1997.

[14] T. J. Endres, B. D. O. Anderson, C. R. Johnson, Jr., and L. Tong, "On the robustness of FIR channel identification from fractionally-spaced received signal second-order-statistics," *IEEE Signal Processing Lett.*, vol. 3, pp. 153–155, May 1996.

[15] T. J. Endres, C. H. Strolle, S. N. Hulyalkar, T. A. Schaffer, A. Shah, M. Gittings, C. Hollowell, A. Bhaskaran, J. Roletter, and B. Paratore, "Carrier independent blind initialization of a DFE using CMA," in *Proc IEEE Signal Process. Workshop Signal Process. Adv. Wireless Commun.*, Annapolis, MD, May 1999, pp. 239–42.

[16] P. Schniter and C. R. Johnson, Jr., "Sufficient conditions for the local convergence of constant modulus algorithms," *IEEE Trans. Signal Processing*, vol. 48, pp. 2785–2796, Oct. 2000.

[17] L. Ljung, *System Identification: Theory for the User*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1999.

[18] I. Fijalkow, C. Manlove, and C. R. Johnson, Jr., "Adaptive fractionally spaced blind CMA adaptation: Excess MSE," *IEEE Trans. Signal Processing*, vol. 46, pp. 227–31, Jan. 1998.

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

Philip Schniter (M'00) was born in Evanston, IL, in 1970. He received the B.S. and M.S. degrees in electrical and computer engineering from the University of Illinois, Urbana-Champaign, in 1992 and 1993, respectively. In 2000, he received the Ph.D. degree in electical engineering from Cornell University, Ithaca, NY.

From 1993 to 1996, he was with Tektronix Inc., Beaverton, OR, as a systems engineer. There, he worked on signal processing aspects of video and communications instrumentation design, including algorithms, software, and hardware architectures. He is currently an Assistant Professor with the Department of Electrical Engineering, The Ohio State University, Columbus. His research interest is signal processing for communication systems, especially blind adaptive equalization.

Dr. Schiter received the 1998 Schlumberger Fellowship and the 1998–1999 Intel Foundation Fellowship while pursuing the Ph.D. degree. He received the 1999 Prize Paper Award from the IEEE Energy Development and Power Generation Committee for work relating to his M.S. thesis.

Raúl A. Casas was born in Denver, CO, in 1972. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from Cornell University, Ithaca, NY, in 1994, 1996, and 1999, respectively.

Since 1999, he has been a systems engineer at NxtWave Communications, Langhorne, PA. His work centers on advanced receiver design for terrestrial digital television. His current research interests include blind equalization and identification of nonlinear systems.

Azzédine Touzni (A'01) was born on July 12, 1969 in Paris, France. He received the Engineering degree in optic telecommunication from the Institut Galilé, Paris, in 1994 and the M.S. degree in signal processing from the University of Cergy-Pontoise (UCP), Cergy, Pointoise, France, in 1995. In 1998, he received the Ph.D degree from UCP.

In 1999, he was with the Institut National de Recherche en Informatique et Automatique (INRIA), Rocquencourt, France, as well as with Cornell University, Ithaca, NY, as a Research Associate with the School of Electrical Engineering. Since June 2000, he has been with NxtWave Communications Inc., Langhorne, PA. His research interests focus on adaptive signal processing applied to digital communications.

C. Richard Johnson, Jr. (F'89) was born in Macon, GA, in 1950. He received the Ph.D. degree in electrical engineering with minors in engineering-economic systems and art history from Stanford University, Stanford, CA, in 1977.

He is currently a Professor of electrical and computer engineering and a member of the Graduate Field of Applied Mathematics at Cornell University, Ithaca, NY. In the past decade, he has held visiting appointments at Stanford University, University of California, Berkeley, Chalmers University of Technology, Göteborg, Sweden, Technical University of Vienna, Vienna, Austria, the National Polytechnic Institute of Grenoble, Grenoble, France, and the Australian National University, Canberra. His research focus over the past decade has been blind adaptive fractionally spaced linear and decision feedback equalization for intersymbol and structured multiuser interference removal. The research of his group at Cornell (http://backhoe.ee.cornell.edu/BERG) is currently supported by the National Science Foundation, Applied Signal Technology, Lucent Technologies—Bell Labs, Aware, and NxtWave Communications.