A Feedback Approach to the Steady-State Performance of Fractionally Spaced Blind Adaptive Equalizers

Junyu Mai and Ali H. Sayed, *Senior Member, IEEE*

***Abstract*—This paper proposes a new approach to the analysis of the steady-state performance of constant modulus algorithms (CMA), which are among the most popular adaptive schemes for blind equalization. A major feature of the proposed feedback approach is that it bypasses the need for working directly with the weight error covariance matrix. In so doing, approximate expressions for the steady-state mean-square error of several CM algorithms are derived, including CMA2-2, CMA1-2, normalized CMA, and a new normalized CMA variant with less bias. A comparison among the various algorithms is also performed, along with several simulation results. The conclusions confirm the superior performance of CMA2-2.**

***Index Terms*—Adaptive filter, blind equalization, constant modulus signal, feedback analysis, mean-square error.**

I. Introduction

AMONG the most popular adaptive schemes for blind equalization are the so-called constant modulus algorithms (CMA's); see [1]–[3] and the many references therein. The update equations of these algorithms are nonlinear in nature, which may explain why only a handful of results are available in the literature regarding their steady-state mean-square-error performance. The difficulty arises from the fact that classical approaches to steady-state performance evaluation often require, as an intermediate step, that a recursion be determined for the covariance matrix of the weight error vector. This step can become a burden for CM algorithms due to their inherent nonlinear updates (see, e.g., the analysis of the constant modulus array algorithm for adaptive beamforming in [4] and the analysis of the performance of CMA for interference cancellation in [5, Sec. 3.3]).

The main objective of this paper is to propose a new approach to the analysis of the steady-state performance of blind adaptive algorithms. A major feature of the approach is that it bypasses the need to work directly with the weight error vector. In so doing, approximate expressions for the steady-state mean-square error of several CM algorithms are derived (including CMA2-2, CMA1-2, normalized CMA, and a new normalized CMA variant with less bias). A comparison among the various algorithms is also performed, along with several simulation results. Our conclusions will further confirm the superior performance of CMA2-2.

The approach in this paper exploits a fundamental energy-preserving relation that, in fact, holds for a general class of adaptive filters and not just CM algorithms [6]. This relation allows us to avoid working directly with the nonlinear update that is characteristic of CM algorithms; it focuses instead on the propagation of error energies through a feedback structure that consists of a lossless feedforward block and a feedback path.

A. Earlier Results in the Literature

Some of the earlier results in the literature on the performance of CM algorithms that are relevant to the discussion in this paper appear in [9]–[13]. The survey article [3] provides a comprehensive list of further additional references on different aspects of CM algorithms. Shynk *et al.* [10] obtain some of the earliest approximations for the mean-square error of the so-called CMA2-2 variant, under the assumption of Gaussian regression vectors. This assumption may not be justified for many communication channels, and the derivation in this paper will provide expressions that result in better approximations for generic regression vectors. Bershad and Roy [11] wrote an early work on the performance of CMA2-2, albeit for a particular class of input signals that are modeled by Rayleigh fading sinusoids. Zeng and Tong [12] studied the mean-square-error of the optimal CM receiver, viz., of the receiver that results by minimizing the CM cost function. The effects of adaptation and gradient noise are not considered. By an ingenious use of Lyapunov stability and averaging analysis, Fijalkow *et al.* [13] obtain an approximate expression for the mean-square error of CMA2-2 that is related to one of our results; though less accurate (see the simulation and comparison results in Section IV-E).

B. Organization of the Paper

The paper is organized as follows. In the next section, we describe the fractionally spaced model adopted in this paper in addition to some of the CM algorithms that we study here. In Section III-B, we motivate and derive the energy-preserving relation and then apply it to CMA2-2. In Sections V and VI, we extend the analysis to CMA1-2 and to normalized CMA. We also develop a normalized CM algorithm with less bias than known normalized variants. Throughout the paper, we provide several

Manuscript received November 20, 1998; revised June 20 1999. This work was supported in part by the National Science Foundation under Awards MIP-9796147 and CCR-9732376. The associate editor coordinating the review of this paper and approving it for publication was Dr. Xiang-Gen Xia.

J. Mai was with the Electrical Engineering Department, University of California, Los Angeles, CA 90024 USA. She is now with the Advanced Research Department, St. Jude Medical Cardiac Rhythm Management Division, Sylmar, CA 91342 USA.

A. H. Sayed is with the Electrical Engineering Department, University of California, Los Angeles, CA 90024 USA (e-mail: sayed@ee.ucla.edu).

Publisher Item Identifier S 1053-587X(00)00099-4.

simulations that compare the theoretical results predicted by our expressions with experimental values. In the concluding section, we compare the various algorithms.

II. THE $T/2$-FRACTIONALLY-SPACED MODEL

Equalization algorithms can be implemented in symbol-spaced form [also called Baud- or T-spaced equalizer form (TSE)] or in fractionally spaced form (FSE). In this paper, we concentrate on fractionally spaced implementaions due to their inherent advantages (see, e.g., [2], [3], [14]–[16]). Thus, consider an FIR channel c of length $2M$ and an FIR equalizer $\overline{w}$ of length $2N$. We split the coefficients of the channel into even- and odd-indexed entries and denote them by

$$\begin{aligned} c_e &\triangleq [c(0), c(2), \cdots, c(2M-2)]^T \\ c_o &\triangleq [c(1), c(3), \cdots, c(2M-1)]^T. \end{aligned}$$

The vectors c_e and c_o are the impulse responses of the two subchannel blocks shown in Fig. 1. In a similar way, we define the two subequalizers

$$\begin{aligned} w_e &\triangleq [\overline{w}(0), \overline{w}(2), \cdots, \overline{w}(2N-2)]^T \\ w_o &\triangleq [\overline{w}(1), \overline{w}(3), \cdots, \overline{w}(2N-1)]^T \end{aligned}$$

which are the impulse responses of the two subequalizer blocks shown in the same figure. The system in Fig. 1 then corresponds to what is called a multichannel model for a $T/2$-fractionally spaced equalizer. This model is well motivated and explained in the survey article [3].

The output $\{y(i)\}$ of the combined channel-equalizer system can be expressed in terms of the transmitted signal $\{s(i)\}$ as follows. Introduce the $(M+N-1)\times N$ prewindowed Toeplitz matrix

$$C_o \triangleq \begin{bmatrix} c(1) & & & \\ c(3) & c(1) & & \\ \vdots & c(3) & \ddots & \\ c(2M-1) & \vdots & \ddots & c(1) \\ & c(2M-1) & \ddots & c(3) \\ & & \ddots & \vdots \\ & & & c(2M-1) \end{bmatrix}$$

and similarly for C_e. Then, define the $(M+N-1)\times 2N$ channel matrix $C = [C_o \;\; C_e]$, the equalizer vector

$$w = \begin{bmatrix} w_e \\ w_o \end{bmatrix}$$

and the input signal vector

$$s_i = [s(i), s(i-1), \cdots, s(i-M-N+2)].$$

Then, $y(i) = s_i C w$. If we further let $u_o(i)$ and $u_e(i)$ denote the input signals to the subequalizers w_e and w_o, respectively, and define the *row* input vectors

$$u_{o,i} = [u_o(i), u_o(i-1), \cdots, u_o(i-N+1)] \quad (1)$$
$$u_{e,i} = [u_e(i), u_e(i-1), \cdots, u_e(i-N+1)] \quad (2)$$

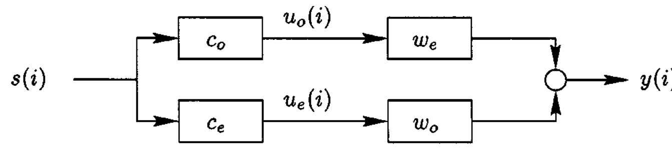

Fig. 1. Multichannel model for $T/2$-FSE.

and

$$u_i = [u_{o,i} \;\; u_{e,i}] \quad (3)$$

then we also have $u_i = s_i C$ and $y(i) = u_i w$.

A. Perfect Equalization

An important result for such fractionally spaced equalizers is the following (see, e.g., [3]). Let $\boldsymbol{c}_e(z)$ and $\boldsymbol{c}_o(z)$ denote the polynomials associated with the even- and odd-indexed subchannels

$$\begin{aligned} \boldsymbol{c}_e(z) &\triangleq c(0) + c(2)z + \cdots + c(2M-2)z^{M-1} \\ \boldsymbol{c}_o(z) &\triangleq c(1) + c(3)z + \cdots + c(2M-1)z^{M-1}. \end{aligned}$$

Then, it can be shown that if these polynomials do not have common zeros, and if $N \geq M-1$, then there exists an equalizer w that leads to an overall channel-equalizer impulse response of the form

$$h_D = e^{j\theta}\text{col}[0, \cdots, 0, 1, 0, \cdots, 0] \quad (4)$$

for some constant phase shift $\theta \in [0, 2\pi]$, and where the unit entry is in some position D, $D \leq M+N-1$. Equalizers w that result in overall impulse responses of the above form are called *zero-forcing* equalizers and will be denoted by w^{ZF}. Thus, under such conditions, the output of the channel-equalizer system will be of the form $y(i) = e^{j\theta}s(i-D)$ for some $\{D, \theta\}$.

The multichannel model of Fig. 1 is the model we are going to study in future sections. For more general T/P-fractionally spaced equalizers, we end up with a similar model with P subchannels and P subequalizers (see, e.g., [15] and [16]), and the results in this paper can be readily extended to this context.

B. Constant Modulus Algorithms

We thus see that under a length-and-zero condition, a finite-length FSE can perfectly equalize a noise-free FIR channel. A blind adaptive equalizer is one that attempts to approximate a zero forcing equalizer without knowledge of the channel impulse response and without direct access to the transmitted sequence $\{s(\cdot)\}$ itself. This is achieved by seeking to minimize certain cost functions that are carefully chosen so that their global minima occur at zero forcing equalizers.

The most popular adaptive blind algorithms are the so-called constant modulus algorithms [17], [18]. They are derived as stochastic gradient methods for minimizing the cost function

$$J_{\text{CM}}(w) = E(|y(i)|^p - R_p|)^2 \quad (5)$$

where w denotes the weight vector to be estimated, and the constant R_p is suitably chosen in order to guarantee that the global minima of $J_{\text{CM}}(w)$ occur at zero forcing solutions (see, e.g., [15], [17], and [19]).

In the next two sections, we study the following two variants: CMA2-2 and CMA1-2. In a later section, we study other variants (known as normalized CM algorithms).[1]

CMA2-2: In this case, we select $p = 2$

$$R_2 = \frac{E|s(i)|^4}{E|s(i)|^2} \tag{6}$$

and the update equation for the weight estimates is given by

$$w_i = w_{i-1} + \mu u_i^* \, y(i) \, \left[R_2 - |y(i)|^2\right] \tag{7}$$

with a step-size μ and where now, $y(i) = u_i w_{i-1}$ is the output of the adaptive equalizer. Here, the symbol $*$ denotes complex conjugate transposition.

CMA1-2: In this case, we select $p = 1$

$$R_1 = \frac{E|s(i)|^2}{E|s(i)|} \tag{8}$$

and the update equation for the weight estimates is given by

$$w_i = w_{i-1} + \mu u_i^* \left[R_1 \, \frac{y(i)}{|y(i)|} - y(i)\right]. \tag{9}$$

Since these algorithms are based on instantaneous approximations of the true gradient vector of the cost function $J_{\rm CM}(w)$, the equalizer output $y(i)$ need not converge to a zero forcing solution of the form $s(i-D)e^{j\theta}$ due to the presence of gradient noise. In the following sections, we derive expressions for the steady-state mean-square error

$$\lim_{i\to\infty} E\left|y(i) - s(i-D)e^{j\theta}\right|^2$$

for adaptive algorithms of the CM class.

III. A New Approach for Steady-State Analysis

As mentioned in the introduction, and as can be seen from the above equations, the updates for CM algorithms are nonlinear in the weight estimates w_i. This may explain why only a few results are available in the literature regarding the steady-state performance of this class of algorithms. The difficulty is because for other adaptive schemes (e.g., of the LMS family), it is common to compute steady-state results by first determining recursions for the squared weight error energy $\|\tilde{w}_i\|^2$ measured relative to some zero-forcing solution, say, $\tilde{w}_i = w^{ZF} - w_i$ (see, e.g., [20]–[23]). This step is a burden for CM algorithms as well as for several other adaptive schemes, due to their nonlinear updates.

Our objective is to propose a new approach for evaluating the steady-state mean-square error of CM algorithms without requiring explicit expressions or recursions for $\|\tilde{w}_i\|^2$. We motivate our approach by first explaining the conventional method for evaluating the mean-square error and by showing the difficulty it encounters when dealing with adaptive filters with nonlinear updates.

A. The Mean-Square Error

Let w^{ZF} denote the zero forcing solution that gives $u_i w^{ZF} = s(i-D)e^{j\theta}$ for some $\{D, \theta\}$. This is guaranteed to exist under some length-and-zero conditions. Now, due to gradient noise, the adaptive equalizer will yield an output $y(i)$ that is distinct from $u_i w^{ZF}$. Let $e_a(i)$ denote the resulting (*a priori*) estimation error as

$$e_a(i) = s(i-D)e^{j\theta} - y(i) = u_i w^{ZF} - u_i w_{i-1} = u_i \tilde{w}_{i-1}.$$

One measure of filter performance is the steady-state mean-square error (MSE)

$$\text{MSE} = \lim_{i\to\infty} E|e_a(i)|^2$$

which is clearly dependent on $\tilde{w}_{i-1}$. It is common in the literature to evaluate this MSE as follows. We first assume that the regression vector u_i is independent of $\tilde{w}_{i-1}$.[2] Then, under this assumption, the above expression for the MSE becomes

$$\text{MSE} = \lim_{i\to\infty} \text{Trace}\,(RF_{i-1}) \tag{10}$$

where $F_i = E\tilde{w}_i\tilde{w}_i^*$ and, assuming stationarity, $R = Eu_i^* u_i$.[3] It is thus customary to determine the steady-state MSE by first determining the steady-state mean-square deviation (MSD) defined by

$$\text{Trace}\,(F) \equiv \lim_{i\to\infty} \text{Trace}\,(F_i) = \lim_{i\to\infty} E\|\tilde{w}_i\|^2. \tag{11}$$

This method of evaluation can become a burden for adaptive algorithms that involve nonlinear updates in $\tilde{w}_i$, as is the case with blind adaptive algorithms. We now describe a new approach for evaluating $E|e_a(\infty)|^2$ that bypasses the need for studying F_i and its limit.

B. A Fundamental Energy-Preserving Relation

The approach is based on a fundamental energy-preserving relation [cf. (20) further ahead], which actually holds for very general adaptive schemes and not just CM algorithms, as explained in [6]. This energy relation was noted and exploited by Sayed and Rupp in [26]–[29] in studies on the robustness and l_2-stability of adaptive filters from a deterministic point of view (see [29]). We review this result below and prepare the notation for later sections.

Consider a general stochastic algorithm of the form

$$w_i = w_{i-1} + \mu u_i^* e_o(i) \tag{12}$$

where $e_o(i)$ denotes an instantaneous error, and u_i a nonzero (row) regression vector. CM algorithms are a special case of the

[1]In our notation, we use parenthesis to refer to scalar variables, e.g., $s(i)$ or $y(i)$ and subscripts to refer to vector quantities, e.g., w_i or u_i. This convention helps distinguish between scalar and vector quantities.

[2]We are not going to impose this condition in our derivation. We are simply using it here to demonstrate the common approach in the literature. We may add that although not true in general, especially for tapped-delay adaptive filter structures, this condition is actually a part of certain widely used independence assumptions in adaptive filter theory [20]. It was shown in [24] and [25], for instance, that for LMS-type scenarios, and for sufficiently small step- sizes, the conclusions that can be obtained from such independence assumptions tend to match reasonably well the real filter performance.

[3]Since we assume in this paper that the input vector u_i is a row vector rather than a column vector, its covariance matrix is therefore defined as $Eu_i^* u_i$ rather than $Eu_i u_i^*$. Our convention of a row vector u_i generally simplifies the notation and avoids an overburden of conjugation symbols.

above for different choices of the function $e_o(i)$. Now, subtract both sides of (12) from some vector w^{ZF} to get the weight error equation

$$\tilde{w}_i = \tilde{w}_{i-1} - \mu u_i^* e_o(i) \tag{13}$$

where $\tilde{w}_i = w^{ZF} - w_i$. Define the *a priori* and *a posteriori* estimation errors $e_a(i) = u_i \tilde{w}_{i-1}$ and $e_p(i) = u_i \tilde{w}_i$. We now show how to rewrite (13) in terms of the error measures $\{\tilde{w}_i, \tilde{w}_{i-1}, e_a(i), e_p(i)\}$ alone. For this purpose, we note that if we multiply (13) by u_i from the left, we obtain

$$e_p(i) = e_a(i) - \mu \|u_i\|^2 e_o(i). \tag{14}$$

Solving for $e_o(i)$ gives

$$e_o(i) = \frac{e_a(i) - e_p(i)}{\mu \|u_i\|^2} \tag{15}$$

so that we can rewrite (13) as

$$\tilde{w}_i = \tilde{w}_{i-1} - \frac{u_i^*}{\|u_i\|^2}[e_a(i) - e_p(i)]. \tag{16}$$

Rearranging (16) leads to

$$\tilde{w}_i + \frac{u_i^*}{\|u_i\|^2} e_a(i) = \tilde{w}_{i-1} + \frac{u_i^*}{\|u_i\|^2} e_p(i). \tag{17}$$

If we define

$$\overline{\mu}(i) \triangleq \frac{1}{\|u_i\|^2} \tag{18}$$

then by squaring (17), we observe that the following energy relation is obtained:

$$\|\tilde{w}_i\|^2 + \overline{\mu}(i)|e_a(i)|^2 = \|\tilde{w}_{i-1}\|^2 + \overline{\mu}(i)|e_p(i)|^2. \tag{19}$$

Interestingly enough, this relation can be obtained by simply replacing the terms of (17) by their respective energies; the cross terms cancel out!. We state this result in the form of a theorem for later reference.

Theorem 1—Energy Relation [26], [27]: Given a generic adaptive algorithm of the form (12), it always holds that

$$\|\tilde{w}_i\|^2 + \overline{\mu}(i)|e_a(i)|^2 = \|\tilde{w}_{i-1}\|^2 + \overline{\mu}(i)|e_p(i)|^2 \tag{20}$$

where $\overline{\mu}(i) = 1/\|u_i\|^2$. □

Relation (20) holds for *any* adaptive algorithm of the form (12); it relates the energies of the weight error vectors at two successive time instants with the energies of the *a priori* and *a posteriori* errors. *No approximations are involved in deriving* (20). The relation also has an interesting physical interpretation. It establishes that the mapping from the variables $\{\tilde{w}_{i-1}, \sqrt{\overline{\mu}(i)}e_p(i)\}$ to the variables $\{\tilde{w}_i, \sqrt{\overline{\mu}(i)}e_a(i)\}$ is energy preserving. Combining (20) with (14), we see that both relations establish the existence of the feedback configuration shown in Fig. 2, where $\mathcal{T}$ denotes the lossless map from $\{\tilde{w}_{i-1}, \sqrt{\overline{\mu}(i)}e_p(i)\}$ to $\{\tilde{w}_i, \sqrt{\overline{\mu}(i)}e_a(i)\}$, and where q^{-1} denotes the unit delay operator. Thus, relation (20) characterizes the energy-preserving property of the feedforward path, whereas relation (14) characterizes the feedback path.

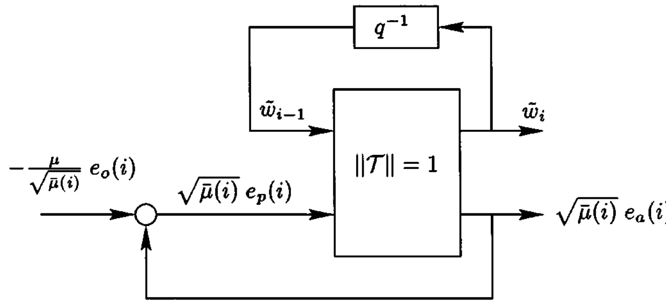

Fig. 2. Lossless mapping and a feedback loop.

C. Significance to MSE Evaluation

We now explain the relevance of the energy relation (20) in the context of MSE evaluation for CM algorithms. (Applications to other classes of adaptive algorithms, in addition to tracking analyzes, are given in [6]–[8].) By taking expectations of both sides of (20), we get

$$E\|\tilde{w}_i\|^2 + E\overline{\mu}(i)|e_a(i)|^2 = E\|\tilde{w}_{i-1}\|^2 + E\overline{\mu}(i)|e_p(i)|^2. \tag{21}$$

Now, recall that our objective in this paper is to evaluate the MSE of CM algorithms in steady state. We are *not* studying conditions under which an algorithm will tend to steady state, which is a separate and complex issue (especially for nonlinear and time-variant filters). Instead, we want to evaluate what performance to expect from an algorithm if it reaches steady state. The convergence to steady state (and, hence, stability) can be studied by relying on results from averaging analysis and from so-called ODE methods (e.g., [30]–[32]); these techniques provide tools that allow one to ascertain, under certain conditions on the data, that there exist small enough step sizes μ for which a filter reaches steady state (see, e.g., [13]).

Thus, assuming filter operation in steady state, we can write

$$E\|\tilde{w}_i\|^2 = E\|\tilde{w}_{i-1}\|^2 \qquad \text{for } i \to \infty. \tag{22}$$

[Similar considerations are also common in the steady-state analysis of other classes of adaptive algorithms (see, e.g., [33]).]

Now, with (22), the effect of the weight error vector is canceled out from (21), and we are reduced to studying only the equality

$$E\overline{\mu}(i)|e_a(i)|^2 = E\overline{\mu}(i)|e_p(i)|^2.$$

This equation provides a relation involving only the desired unknown $e_a(i)$ since $e_p(i)$ is itself a function of $e_a(i)$, as evidenced by (14). Thus, by solving the above equation as $i \to \infty$, we can obtain an expression for the MSE.

Theorem 2—Identity for MSE Analysis: Consider a generic adaptive algorithm of the form (12). In steady state (as $i \to \infty$), when (22) holds, we obtain

$$E\overline{\mu}(i)|e_a(i)|^2 = E\overline{\mu}(i)\left|e_a(i) - \frac{\mu}{\overline{\mu}(i)}\, e_o(i)\right|^2. \tag{23}$$

□

IV. Steady-State Analysis of CMA2-2

We now demonstrate how the result of Theorem 2, which holds for generic adaptive schemes of the form (12), can be applied to the CMA2-2 recursion (7). In later sections, we consider other CM algorithms.

The derivation in the sequel relies on some statistical assumptions (four in total), the introduction of which simplifies the analysis. Although these assumptions may not hold in general, they are realistic for sufficiently small step sizes and, as we shall see from several simulations, lead to good fits between our theoretical results and the simulation results.[4] Following each assumption, we will provide a brief motivation and justification for its use.

A. Two Initial Assumptions

The analysis that follows for CMA2-2 is based on the following two assumptions in *steady-state* ($i \to \infty$).

Assumption I.1: The transmitted signal $s(i-D)$ and the estimation error $e_a(i)$ are independent in steady state so that $Es^*(i-D)e_a(i) = 0$ since $s(i-D)$ is assumed zero mean.

This is a reasonable assumption since it essentially requires the estimation error $\{e_a(i)\}$ of the equalizer to be insensitive, in steady-state, to the actual transmitted symbols $\{s(i)\}$. For example, for symbols from a 2-PAM constellation $s(i) \in \{+1, -1\}$, this means that we are requiring the behavior (or distribution) of the error $e_a(i)$, after the equalizer has converged to steady state, to be insensitive to whether the polarity of $s(i)$ is $+1$ or -1.

Assumption I.1 can be replaced by the following two conditions, which also enable us to conclude that $Es^*(i-D)e_a(i) = 0$.

i) In steady state, CMA2-2 converges in the mean to a zero forcing solution, i.e., the mean of the combined channel-equalizer response $h_i = Cw_i$ converges to $h_D = e^{j\theta}$ col$[0, \cdots 0, 1, 0, \cdots 0]$ for some $\{D, \theta\}$.

ii) $s(i-D)$ and w_i are independent as $i \to \infty$. That is, in steady state, the equalizer operates independently of the transmitted signals. This is a common assumption for steady-state analysis (see, e.g., [13]).

Assumption I.2: The scaled regressor energy $\mu^2\|u_i\|^2$ is independent of $y(i)$ in steady state.

This assumption requires the scaled energy of the input vector and not the input vector itself to be independent of the equalizer output. The assumption actually becomes realistic for longer filter lengths and for sufficiently small step sizes. To see this, assume the input sequence $\{s(i)\}$ is i.i.d., and note that the variance of the quantity $\|u_i\|^2$ will be of the order of N (the equalizer length).[5] Hence, if the step-size μ is of the order of $1/N$ (or less), then the variance of $\mu^2\|u_i\|^2$ is of the order of $1/N$ (or less), which decreases with increasing filter length. This means that $\mu^2\|u_i\|^2$ will eventually tend to a constant and can, therefore, be assumed to be independent of $y(i)$. Note that by the same argument, we can also assume that $\mu^2\|u_i\|^2$ is independent of $e_a(i)$ in steady state. (We may add that an assumption similar to I.2 is also used in [13].)

B. The Case of Real-Valued Data

We start our analysis with the case of real-valued data $\{s(i), y(i), u_i\}$ (e.g., data from a PAM constellation). In the next section, we consider complex-valued data. It turns out that the expressions for the MSE of CMA2-2 are distinct in both cases, whereas those for CMA1-2 are not.

For real-valued data, the zero forcing response h_D that the adaptive equalizer attempts to achieve [cf. (4)] can be of either form $h_D = \pm[0, \cdots, 0, 1, 0, \cdots, 0]$. In the following, we continue with the choice $h_D = [0, \cdots, 0, 1, 0, \cdots, 0]$, which yields

$$e_a(i) = s(i-D) - y(i).$$

A similar analysis holds for the case $h_D = [0, \cdots, 0, -1, 0, \cdots, 0]$.

Now, the relation (23) in the CMA2-2 context leads to the equality, for $i \to \infty$

$$E\overline{\mu}(i)|e_a(i)|^2 = E\overline{\mu}(i)\left|e_a(i) - \frac{\mu}{\overline{\mu}(i)}\, y(i)(R_2 - |y(i)|^2)\right|^2. \tag{24}$$

We will write more compactly (here and throughout the paper)

$$e_a \triangleq e_a(i), \quad \overline{\mu} \triangleq \overline{\mu}(i), \quad y \triangleq y(i), \quad u \triangleq u_i, \quad s \triangleq s(i-D), \qquad \text{for } i \to \infty$$

so that (24) becomes, after expanding

$$E\overline{\mu}|e_a|^2 = E\overline{\mu}|e_a|^2 - \underbrace{2\mu E e_a y(R_2 - y^2)}_{A} + \underbrace{\mu^2 E\|u\|^2 y^2(R_2 - y^2)^2}_{B}.$$

This implies that the terms A and B should coincide. From this equality, we can obtain an approximate expression for the steady-state MSE $E|e_a|^2$ as we now verify. (In the argument below, we assume that when the adaptive filter reaches steady state, the value of e_a^2 is reasonably small.)

Theorem 3—MSE for Real CMA2-2: Consider the CMA2-2 recursion (7) with real-valued data $\{s(i), u_i, y(i)\}$. Under Assumptions I.1 and I.2, it holds that for sufficiently small μ, the steady-state MSE can be approximated by

$$E|e_a|^2 \approx \mu\,\frac{E(s^2R_2^2 - 2R_2s^4 + s^6)}{2E(3s^2 - R_2)}\, E\|u\|^2. \tag{25}$$

Proof: We first evaluate A. Replacing y by $s - e_a$, we obtain

$$A = 2\mu E(se_aR_2 - e_a^2R_2 - s^3e_a + 3s^2e_a^2 - 3se_a^3 + e_a^4).$$

[4] Similar assumptions are very common in the adaptive filtering literature for FIR structures, where they are collectively known as the independence assumptions. As mentioned in a previous footnote, although the independence assumptions do not hold in general, they still lead to realistic conclusions for sufficiently small step sizes [24], [25], [33].

[5] This is obvious if the individual entries of u_i are i.i.d. Some calculations will show that a similar conclusion holds, in general, when the entries of u_i are taken as the outputs of an FIR channel with i.i.d. input.

Using Assumption I.1 and neglecting $2\mu E e_a^4$ for small μ and small e_a^2 leads to the approximation $A \approx 2\mu E(3s^2 - R_2) \cdot E e_a^2$. We now evaluate B

$$B = E\left(\mu^2 \|u\|^2 (s - e_a)^2 \left[R_2 - (s - e_a)^2\right]^2\right).$$

With Assumption I.2, we can rewrite B as in (25a), shown at the bottom of the page. Again, when μ and e_a^2 are small enough, we can ignore the term C and write

$$B \approx \mu^2 E(s^2 R_2^2 - 2R_2 s^4 + s^6) \cdot E\|u\|^2.$$

From the equality $A = B$, we obtain (25). □

C. The Case of Complex-Valued Data

The expression for the MSE of CMA2-2 in the complex case differs from the one we derived above for the real case, as we shall promptly verify.

In the complex case, as in [17], we study signal constellations that satisfy the circularity condition

$$E s^2(i) = 0 \tag{26}$$

in addition to the condition $E(2|s(i)|^2 - R_2) > 0$, which holds for most constellations.

Theorem 4—MSE For Complex CMA2-2: Consider the CMA2-2 recursion (7), and assume complex-valued data $\{s(i), y(i), u_i\}$ satisfying (26). Under Assumptions I.1 and I.2, and for sufficiently small μ, the steady-state MSE can be approximated by

$$E|e_a|^2 \approx \mu \frac{E(|s|^2 R_2^2 - 2R_2|s|^4 + |s|^6)}{2E(2|s|^2 - R_2)} E\|u\|^2. \tag{27}$$

Proof: Starting with (23), we now obtain

$$\begin{aligned} E\overline{\mu}|e_a|^2 = & E\overline{\mu}|e_a|^2 \\ & - \underbrace{\mu E e_a^* y(R_2 - |y|^2) + \mu E e_a y^*(R_2 - |y|^2)}_{D} \\ & + \underbrace{\mu^2 E\|u\|^2 |y|^2 (R_2 - |y|^2)^2}_{F}. \end{aligned}$$

Substituting y by $se^{j\theta} - e_a$, we get

$$\begin{aligned} D = & \mu R_2 E\left(se^{j\theta} e_a^* + s^* e^{-j\theta} e_a\right) \\ & - \mu E\left(|s|^2 s e^{j\theta} e_a^* + |s|^2 s^* e^{-j\theta} e_a\right) \\ & - 3\mu E\left(|e_a|^2 s^* e^{-j\theta} e_a + |e_a|^2 s e^{j\theta} e_a^*\right) \\ & - 2\mu E\left(R_2|e_a|^2 + 4|s|^2|e_a|^2 + 2|e_a|^4\right) \\ & + \mu E\left(s^{*2} e^{-2j\theta} e_a^2 + s^2 e^{2j\theta} e_a^{*2}\right). \end{aligned}$$

By using (26) and Assumption I.1, the term D can be simplified to $D \approx 2\mu E(2|s|^2 - R_2) \cdot E|e_a|^2$. Similarly, expanding F and using the same approximations as in the real-valued case, we obtain

$$F \approx \mu^2 E(R_2^2 |s|^2 - 2R_2|s|^4 + |s|^6) \cdot E\|u\|^2.$$

Then, from $D = F$, we get (27). Note that (27) will not be negative because of

$$E(|s|^2 R_2^2 - 2R_2|s|^4 + |s|^6) = E|s|^2 (R_2 - |s|^2)^2 \geq 0$$

and $E(2|s|^2 - R_2) > 0$. □

Comparing the results we get for the real-valued and complex-valued cases, we see that they are similar except for a coefficient in the denominator expressions (in the real case it is equal to 3 and in the complex case it is equal to 2). Moreover, some useful conclusions can be drawn from these results.

1) The steady-state MSE of CMA2-2 is linearly proportional to the step-size μ and to the received signal variance $E\|u\|^2$, which agrees with the asymptotic MSE result for the symbol-spaced (TSE) CM algorithm in [10] and [34]. This property is also similar to that of LMS.
2) For constant modulus signals $\{s(\cdot)\}$, we get $R_2 = 1$. According to (25) and (27), we then obtain $E|e_a|^2 = 0$. This is also the same as LMS in the absence of noise.
3) For nonconstant modulus signals, the MSE will not be zero, even when there is no system noise. This is because the instantaneous error $e_o(i)$ for CMA2-2 will be nonzero, even when $y(i) = s(i - D)e^{j\theta}$. The equalizer weight vector w_i keeps updating itself by a nonvanishing term and jitters around the mean solution. This property is different from LMS, where the instantaneous error will be equal to zero when the system is perfectly equalized.

D. Simulation Results for CMA2-2

Before proceeding to other CM algorithms, we provide some simulation results that compare the experimental performance with the one predicted by the previous theorems. The simulations will show than the theoretical values predicted by the expressions in Theorems 3 and 4 match reasonably well the experimental results. The channel considered in this simulation is given by $c = [0.1, 0.3, 1, -0.1, 0.5, 0.2]$. A four-tap FIR filter is used as a $T/2$-fractionally spaced equalizer.

1) Constant Modulus Signals: A computer simulation was first done for real and constant modulus signals, i.e., for binary data. With a step-size $\mu = 0.01$, after 10 000 iterations, CMA2-2 was observed to converge to a zero forcing solution with MSE as low as –120 dB, i.e. MSE $= 10^{-12}$, which can be considered zero. This result agrees with our analytical result that the MSE for constant modulus signals is zero.

$$B = \mu^2 E(s^2 R_2^2 - 2R_2 s^4 + s^6) \cdot E\|u\|^2 + \underbrace{\mu^2 E\left[(R_2^2 - 12R_2 s^2 + 9s^4)e_a^2 + 15 s^2 e_a^4 + e_a^6 - 2R_2 e_a^4\right] \cdot E\|u\|^2}_{C} \tag{25a}$$

TABLE I
MSE OF CMA2-2 VERSUS STEP-SIZE FOR 6-PAM SIGNALS

step-size μ	1×10^{-6}	1×10^{-5}	2×10^{-5}	5×10^{-5}	1×10^{-4}
experimental MSE (dB)	33.0	21.9	19.4	15.3	12.1
MSE from Thm. 3 (dB)	31.2	21.2	18.1	14.7	11.2
MSE from [13] (dB)	36.3	26.3	23.2	19.8	16.3

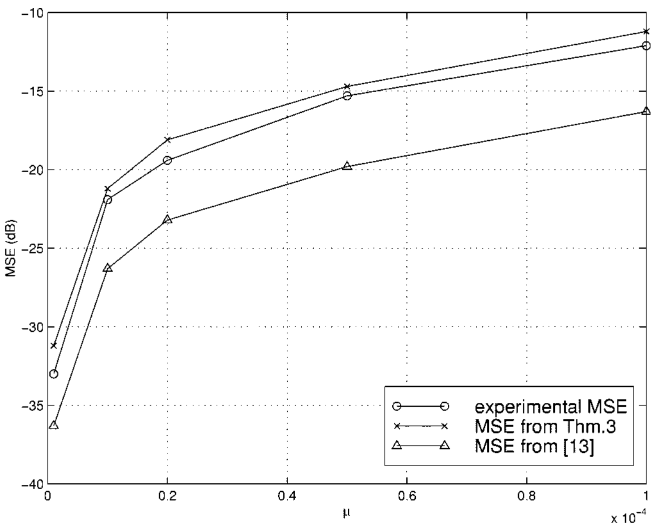

Fig. 3. Experimental and theoretical curves for the steady-state MSE as a function of the step-size, for CMA2-2 with input signals from a 6-PAM constellation.

2) Real and Nonconstant Modulus Signals: In this simulation, the transmitted signal was 6-PAM constellated $s(i) \in \{5, 3, 1, -1, -3, -5\}$ with $E|s(i)|^6 = 5451.7$, $E|s(i)|^4 = 235.7$, $E|s(i)|^2 = 11.67$, and $R_2 = 20.2$. The value of $||u_i||^2$ is the norm of the received signal vector. The value of $E||u_i||^2$ was computed as the average over 3000 realizations of $||u_i||^2$. The first two lines of Table I show the experimental MSE and the theoretical MSE from Theorem 3, where the value of experimental MSE was obtained as the average over 20 repeated experiments. Fig. 3 is a plot of the experimental MSE and the theoretical MSE versus the step-size μ (it also contains one more MSE curve to be discussed in Section IV-E).

3) Complex and Nonconstant Modulus Signals: With the same channel and equalizer, we obtained the MSE for 16-QAM signals. Now, $E|s(i)|^6 = 1960$, $E|s(i)|^4 = 132$, $E|s(i)|^2 = 10$, and $R_2 = 13.2$. The results are shown in Table II and Fig. 4.

E. Comparison with Related Results in the Literature

As mentioned in the introduction, an approximate expression for the MSE of CMA2-2 was also derived in [13]. The derivation assumed real-valued data and that $E||u_{e,i}||^2 = E||u_{o,i}||^2$ for $T/2$-fractionally spaced equalization [recall the definition of $\{u_{e,i}, u_{o,i}\}$ from (1)–(3)]. It further led to the result

$$E|e_a|^2 = \mu\,\frac{\dfrac{E|s|^6}{(E|s|^2)^3} - \rho^2}{2(3-\rho)} \cdot (E|s|^2)^2 \cdot 2E||u_e||^2 \qquad (28)$$

TABLE II
MSE OF CMA2-2 VERSUS STEP-SIZE FOR 16-QAM SIGNALS

step-size μ	1×10^{-6}	2×10^{-6}	5×10^{-6}	1×10^{-5}	2×10^{-5}
experimental MSE (dB)	35.9	31.7	27.9	24.5	21.7
MSE from Thm. 4 (dB)	33.6	30.6	26.6	23.6	20.6

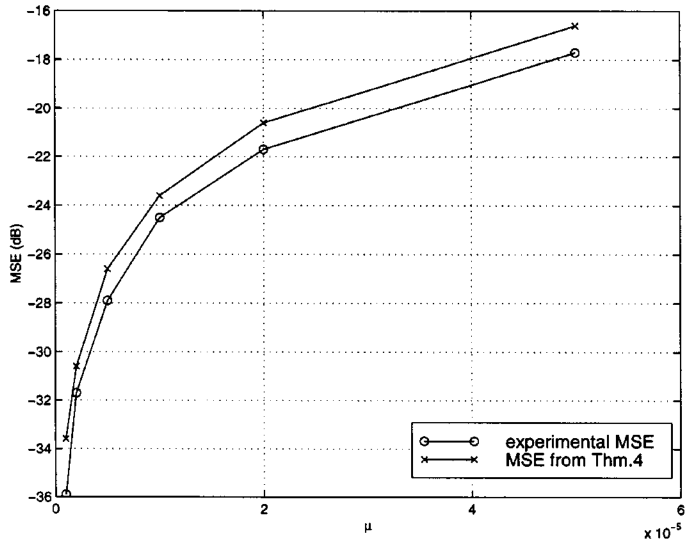

Fig. 4. Experimental and theoretical curves for the steady-state MSE as a function of the step-size for CMA2-2 with input signals from a 16-QAM constellation.

where $\rho = E|s|^4/(E|s|^2)^2$, but since we can write

$$\frac{\dfrac{E|s|^6}{(E|s|^2)^3} - \rho^2}{2(3-\rho)} \cdot (E|s|^2)^2 = \frac{E(|s|^2R_2^2 - 2R_2|s|^4 + |s|^6)}{2E(3|s|^2 - R_2)}$$

we see that the result in (28) actually coincides with our result for real-valued data (cf. Theorem 3), except that the term $E||u||^2$ in our expression is replaced by the term $2E||u_e||^2$ in the above expression from [13]. In other words, the result of [13] assumes that the average energy of all input vectors across the subequalizers are identical, i.e., for $T/2-$equalizers, $E||u_o||^2 = E||u_e||^2$. When the input energy across the subequalizers is not uniform, both expressions will, of course, be different.

Table I and Fig. 3 compare the experimental MSE with (28) and our result (for the real-valued case since [13] considered this case only). Our results seem to be more accurate in part because the input energy across subequalizers is not uniform in general. We may further remark that the approach in [13], although complementary, is considerably different from the approach of this paper. The authors of [13] employ averaging theory [32], solve a Lyapunov equation to find $E||\tilde{w}_\infty||^2$, and then calculate $E|e_a(\infty)|^2$. Here, we started from the generic equality $E\overline{\mu}|e_a|^2 = E\overline{\mu}|e_p|^2$ and solved directly for $E|e_a|^2$. In the next sections, we further extend this approach to other kinds of CM algorithms.

In earlier work [10], an approximate expression was also obtained for the MSE of CMA2-2. However, as mentioned earlier, the analysis in this reference assumes baud-spaced equalizers and Gaussian regression vectors.

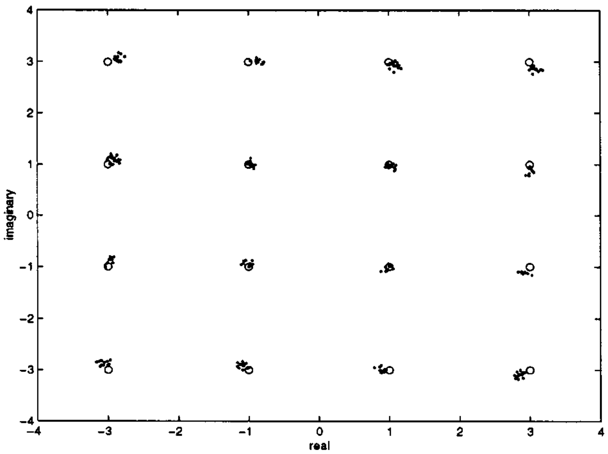

Fig. 5. Typical plot of transmitted signals $s(\cdot)$ (denoted by "o") and equalizer outputs $y(\cdot)$ (denoted by "·") for a 16-QAM constellation.

V. Steady-State Analysis of CMA1-2

We now extend the earlier results to the CMA1-2 recursion (9). In this case, the expressions for the MSE for both real and complex-valued data will coincide. For this reason, we shall consider only the complex-valued case.

A. Two More Assumptions

In addition to Assumptions I.1 and I.2, we need the following two assumptions (also in steady-state).

Assumption I.3: The output $y(i)$ of the equalizer is distributed symmetrically around the transmitted signal $s(i-D)$ in steady state so that $E|y(i)| = E|s(i-D)|$.

Fig. 5 is a plot of the steady-state output $y(\cdot)$ (which is denoted by "·") and the transmitted signal $s(\cdot)$ (which is denoted by "o") in one simulation for a 16-QAM data constellation. We see that we can reasonably assume that the expected value of $|y(i)|$ is equal to the expected value of $|s(i-D)|$.

Assumption I.4: The *a priori* error $e_a(i)$ is independent of $\operatorname{sign} y(i)$ in steady-state, and $E \operatorname{sign} y(i) = 0$ so that $Ee_a(i) \operatorname{sign} y(i) = 0$.

This assumption is again reasonable in steady state and for sufficiently small step sizes. This is because in this situation, we obtain relatively small estimation errors $e_a(i)$ so that the sign of $y(i)$ is essentially determined by the sign of $s(i-D)e^{j\theta}$, which, as explained in Assumption I.2, can be taken to be independent of $e_a(i)$. [We should mention that for complex-valued data, we define $\operatorname{sign} y(i) = (y(i)/|y(i)|)$.]

B. The Case of Complex-Valued Data

Returning to the CMA1-2 recursion (9), we see that the relation (23) between $e_p(i)$ and $e_a(i)$ reduces to

$$e_p(i) = e_a(i) - \mu \|u_i\|^2 [R_1 \operatorname{sign} y(i) - y(i)]. \tag{29}$$

Starting again with the basic equation (23) and using (29), we obtain in steady-state

$$E\overline{\mu}|e_a|^2 = E\overline{\mu} \cdot \left| e_a - \frac{\mu}{\overline{\mu}}[R_1 \operatorname{sign}(y) - y] \right|^2. \tag{30}$$

This equation can be used to establish the following result.

TABLE III
MSE of CMA1-2 versus Step-Size for 6-PAM Signals

step-size μ	5×10^{-5}	1×10^{-4}	2×10^{-4}	5×10^{-4}	1×10^{-3}
experimental MSE (dB)	28.0	23.0	20.8	15.9	13.8
MSE from Thm. 5 (dB)	25.7	22.6	19.6	15.5	12.6

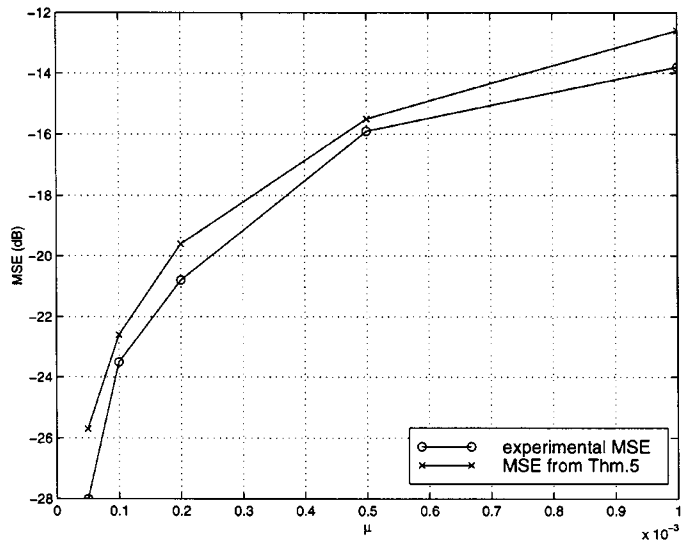

Fig. 6. Experimental and theoretical curves for the steady-state MSE as a function of the step size for CMA1-2 with input signals from a 6-PAM constellation.

TABLE IV
MSE of CMA1-2 versus Step-Size for 16-QAM Signals

step-size μ	1×10^{-5}	2×10^{-5}	5×10^{-5}	1×10^{-4}	2×10^{-4}
experimental MSE (dB)	35.5	33.2	29.0	26.8	23.7
MSE from Thm. 5 (dB)	37.9	34.9	30.9	27.9	24.9

Theorem 5—MSE for Complex CMA1-2: Consider the recursion (9) for complex-valued data and Assumptions I.1–I.4. It then holds, for sufficiently small μ, that the steady-state MSE can be approximated by

$$E|e_a|^2 \approx \frac{\mu}{2}\left(R_1^2 + E|s|^2 - 2R_1 E|s|\right) \cdot E\|u\|^2. \tag{31}$$

Proof: See Appendix A where, as in the CMA2-2 case, we again invoke the fact that e_a^2 is small in steady state. □

C. Simulation Results

We employ the same channel as in the CMA2-2 case. For real-valued signals, we used a 6-PAM data constellation. Table III and Fig. 6 showsthe experimental and theoretical values of the MSE for 6-PAM. Table IV and Fig. 7 show the same values for 16-QAM signals.

From the above simulations, we can see that the theoretical results matchμ reasonably well the experimental results. The MSE of CMA1-2 is also seen to be proportional to .

VI. Normalized CM Algorithms

The normalized CM algorithm has been motivated by the desire to speed up the convergence of CMA1-2 [35], [36]. This, however, leads to a biased estimator $y(i)$ for the transmitted

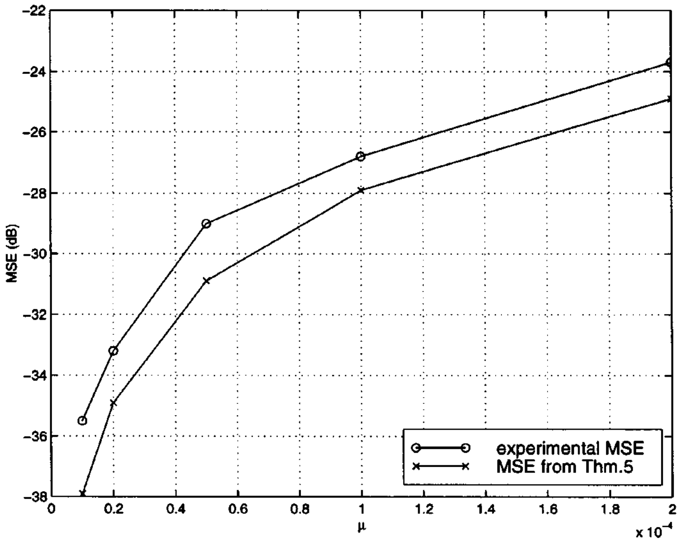

Fig. 7. Experimental and theoretical curves for the steady-state MSE as a function of the step size for CMA1-2 with input signals from a 16-QAM constellation.

signal $s(i-D)$ when the signal constellation is not constant modulus. In this section, we introduce a variant that leads to less bias than earlier algorithms.

The recursion for normalized CMA has the general form

$$w_i = w_{i-1} + \frac{\mu}{||u_i||^2} u_i^* [R \,\text{sign}\, [y(i)] - y(i)] \qquad (32)$$

where R is a design parameter. Compared with the CMA1-2 recursion (9), we see that the constant step size of CMA1-2 is now replaced by a time-variant step size $\mu(i) = (\mu/||u_i||^2)$.

For constant modulus signals $\{s(\cdot)\}$, the most reasonable selection for R is the magnitude of the transmitted signals $R = |s(i)|$. For nonconstant modulus signals, on the other hand, we need to choose other values for R. In [37], it was suggested that we choose, for any $p \geq 1$

$$R = \frac{E|s(i)|^{2p}}{E|s(i)|^p}.$$

For $p=1$, this leads to the choice

$$R_1 = \frac{E|s(i)|^2}{E|s(i)|} \qquad (33)$$

and for $p=2$

$$R_2 = \frac{E|s(i)|^4}{E|s(i)|^2}. \qquad (34)$$

For example, 4-PAM signals $s(i) \in \{-3, -1, 1, 3\}$, we obtain

$$R_1 = \frac{E|s(i)|^2}{E|s(i)|} = 2.5, \quad R_2 = \frac{E|s(i)|^4}{E|s(i)|^2} = 8.2.$$

Fig. 8 demonstrates the bias problem that arises when the normalized CMA recursion is used with the above choices for R, viz., R_1 and R_2. The two left-most plots in the figure show the equalizer outputs with 4-PAM inputs $s(i) \in \{-3, -1, 1, 3\}$ for R_1 and R_2. (The right-most plot uses a different value for R, which we shall derive further ahead.) The channel was $c = [0.1, 0.3, 1, -0.1, 0.5, 0.2]$, and we implemented a four-tap $T/2$ fractionally spaced equalizer. We can see that both plots on the left lead to biased steady state solutions. For example, when $R = 2.5$, $y(i) \in \{4, 1.2, -1.2, -4\}$, on average.

We now propose to select R differently by minimizing the steady-state MSE relative to a zero forcing solution. We focus here on real-valued data. Using the normalized CM recursion (32) and relation (23), we find that

$$e_p(i) = e_a(i) - \mu[R \,\text{sign}\, y(i) - y(i)] \qquad (35)$$

so that (23), as $i \to \infty$, reduces to

$$E\overline{\mu}|e_a|^2 = E\overline{\mu}\,|e_a - \mu(R \,\text{sign}\, y - y)|^2. \qquad (36)$$

As before, we can proceed to evaluate $E|e_a|^2$. However, our earlier derivations were all based on Assumption I.1 and, because of the bias problem, this assumption is no longer satisfied by the normalized CM algorithm for the above values of R (R_1 and R_2).

Note, however, that the larger the bias the larger the value of the steady-state MSE. This suggests selecting R by minimizing the MSE. Such a value for R would result in reduced bias, in which case, we could assume that Assumption I.1 is enforced at least approximately (as is demonstrated by the right-most plot of Fig. 8 for the value of R we will obtain).

In this case, and using Assumptions I.1I.4, we can establish that for sufficiently small and e_a, the resulting steady-state MSE would be (see Appendix B)

$$E|e_a|^2 \approx \frac{\mu}{2}\left(R^2 + E|s|^2 - 2RE|s|\right). \qquad (37)$$

We can now seek that value for R that minimizes (37). Setting the derivative of (37) with respect to R equal to zero leads to the choice $R_{opt} = E|s|$, and the corresponding MSE will be $Ee_a^2 = \mu(E|s|^2 - (E|s|)^2)$. Therefore, with $R = E|s|$, we obtain the variant

$$w_i = w_{i-1} + \frac{u_i^*}{||u_i||^2}\,[E|s(i)|\text{sign}\, y(i) - y(i)]. \qquad (38)$$

The simulation result in Fig. 8 shows that this selection for R leads to a considerably smaller offset and MSE.

Moreover, Table V and Fig. 9 show the values of experimental MSE and theoretical MSE for different step-sizes for 6-PAM signals using (38). We can see that the theoretical MSE does not match closely the experimental results. The reason is that our selection for R_{opt}, although close, does not fully result in unbiased estimation. Thus, the bias problem makes it difficult to satisfy Assumption I.1.

Finally, as mentioned in the introduction of this section, normalized CM algorithms are motivated by the desire to speed up the convergence of CMA1-2. In Fig. 10, we compare the convergence rate of both these algorithms by using the above choice for R, $R_{opt} = E|s|$. The channel is $c = [-0.0901, 0.6853, 0.7170, 0.0901]$, and the equalizer is a two-tap FIR filter. The input constellation is 4-PAM. We use the step-size $\mu = 0.002$ for both algorithms. Unlike the case of

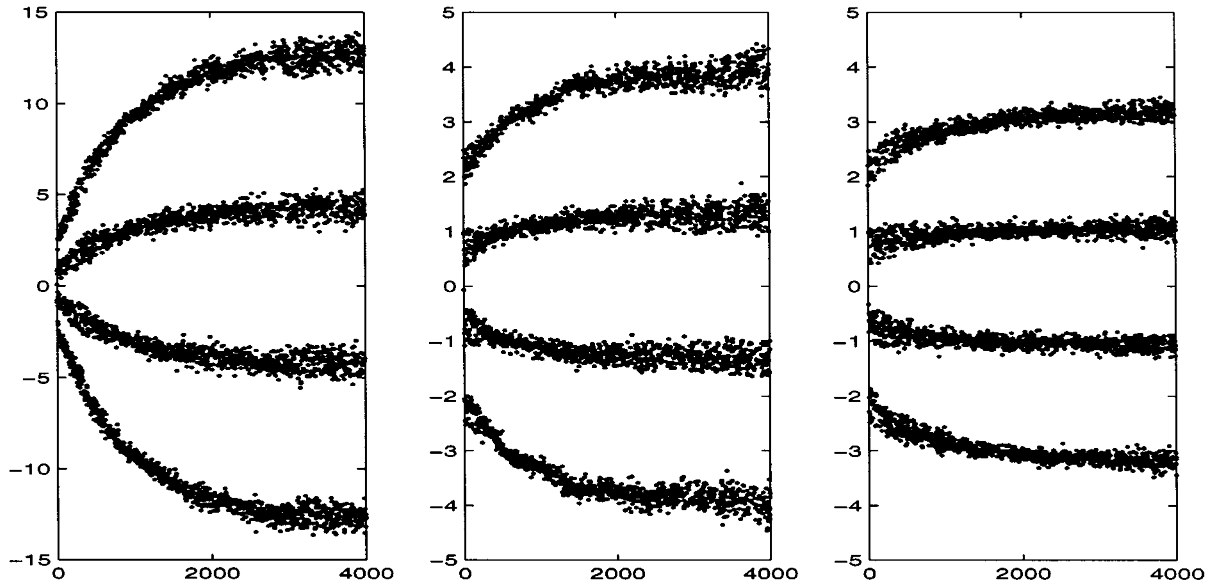

Fig. 8. Equalizer outputs $y(i)$ of normalized CMA for three different choices of the parameter R. (left) $R = (E|s(i)|^4/E|s(i)|^2) = 8.2$. (middle) $R = (E|s(i)|^2/E|s(i)|) = 2.5$. (right) $R = E|s(i)| = 2$. The horizontal axis denotes iteration time while the vertical axis denotes amplitude.

TABLE V
MSE OF NORMALIZED CMA (38) VERSUS STEP-SIZE FOR SIX-PAM SIGNALS

step-size μ	5×10^{-3}	1×10^{-3}	1×10^{-4}	1×10^{-5}	1×10^{-6}
experimental MSE (dB)	18	21	30	48	56
MSE from (37) (dB)	26	33	43	53	63

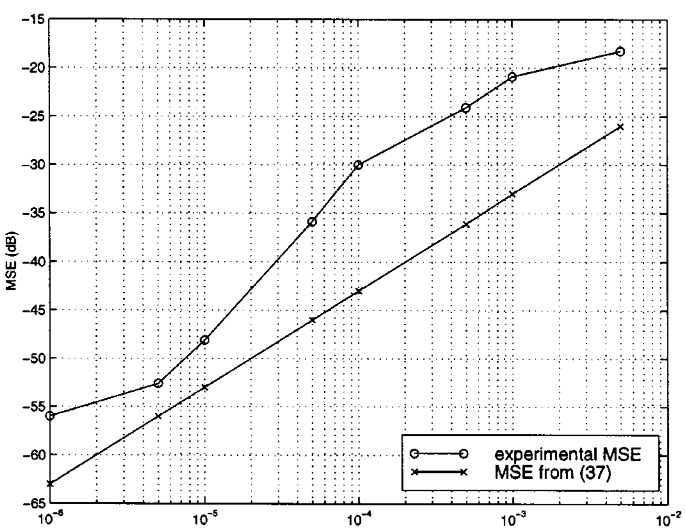

Fig. 9. Experimental and theoretical curves for the steady-state MSE as a function of the step size for normalized CMA (38) with input signals from a 6-PAM constellation.

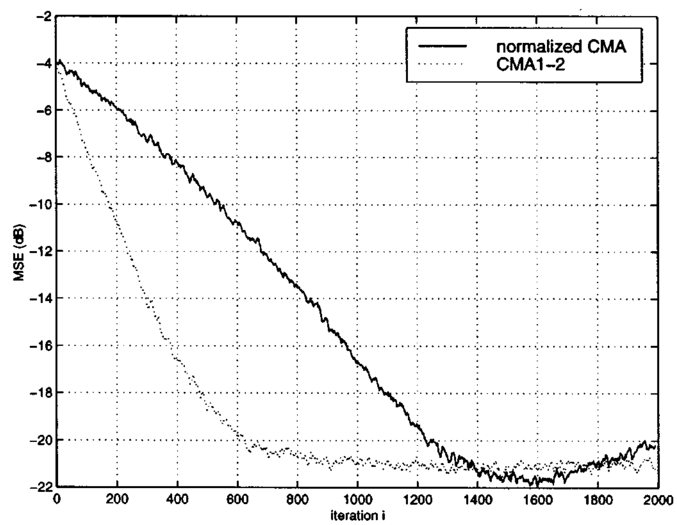

Fig. 10. Comparison of the convergence rates of CMA1-2 and normalized CMA for 4-PAM signals. The figure shows two ensemble-average learning curves obtained by averaging over several experiments.

constant modulus signals, the simulation shows that normalized CMA need not converge faster than CMA1-2. The figure plots ensemble-average curves for the *a priori* estimation error energy $\{|e_a(\cdot)|^2\}$ in decibels, averaged over ten experiments. A similar conclusion holds for the other choices of R in (33) and (34).

VII. CONCLUDING REMARKS

In this paper, we studied the steady-state performance of several blind adaptive algorithms of the constant modulus type, namely, CMA2-2, CMA1-2, and normalized CMA. Analytical expressions for the steady-state mean-square error (MSE) were calculated and verified by computer simulations. From this study, we conclude the following.

1) The fundamental energy-preserving relation described in SectionIII-B is a useful property for the analysis of the steady-state performance of gradient-based adaptive algorithms. By using this relation, we could calculate the MSE of CM algorithms in a simpler way than other methods.

2) For nonconstant modulus signals, the MSE of CMA1-2 and CMA2-2 will not converge to zero even when there is no channel noise. Moreover, the MSE of CMA1-2 and CMA2-2 are determined by the signal constellation size and are proportional to the step size of the algorithms and to the received signal energy (or variance).
3) For constant modulus signals, the MSE of CMA1-2 and CMA2-2 will converge to zero when there is no channel noise. The step size should be sufficiently small to guarantee stable operation of the equalizer for both algorithms. For CMA1-2, because of the existence of undesired local minima, special care is needed with the initial condition.
4) Normalized CMA is a faster algorithm than CMA1-2 for constant-modulus signals. For nonconstant modulus signals, however, normalized CMA will converge to a biased solution. We showed in SectionVI how to decrease the bias by designing a new normalized CM algorithm.
5) Our analysis suggests that CMA2-2 has the best performance among the algorithms we discussed in this paper. When implemented in a fractionally spaced form, it has no undesired minima, it converges faster than CMA1-2, it gives an unbiased solution for both constant and nonconstant modulus signals, and it requires only simple calculations.

APPENDIX A
PROOF OF THEOREM 5

Expanding the right-hand side of (30) leads to /belowdisplayskip16pt

$$E\overline{\mu}|e_a|^2 = E\overline{\mu}|e_a|^2 - \underbrace{\mu[Ee_a^*(R_1 \,\text{sign}\,(y) - y) + e_a(R_1 \,\text{sign}\,(y) - y)^*]}_{J} + \underbrace{\mu^2 E|R_1 \,\text{sign}\,(y) - y)|^2 \cdot ||u||^2}_{K}.$$

Hence, $J = K$. Replacing y by $se^{j\theta} - e_a$ and using Assumptions I.1 and I.4, we get $J = 2\mu E|e_a|^2$. Using Assumption I.2, the term K becomes

$$K = \mu^2 E||u||^2 \cdot E(R_1^2 - 2R_1|y| + |y|^2).$$

Now, from Assumptions I.1 and I.3, we get $E|y| = E|s|$ and $E|y|^2 = E|s|^2 + E|e_a|^2$. Then, $K = \mu^2 E||u||^2 \cdot E(R_1^2 - 2R_1|s| + |s|^2 + |e_a|^2)$. Ignoring the term $\mu^2 E||u||^2 \cdot E|e_a|^2$µ, when and e_a are sufficiently small, we get

$$K \approx \mu^2 E||u||^2 \cdot E(R_1^2 - 2R_1|s| + |s|^2).$$

Using $J = K$, we are led to (31). Note that the expression for the MSE cannot be negative because $R_1^2 + E|s(i)|^2 - 2R_1E|s(i)| = E(R_1 - |s(i)|)^2 \geq 0$. □

APPENDIX B
DERIVATION OF (37)

From (35) and (36), we get /belowdisplayskip8pt

$$\begin{aligned} &E\overline{\mu}(i)|e_a(i)|^2 \\ &= E\left(\overline{\mu}(i)\,|e_a(i) - \mu[R\,\text{sign}(y(i)) - y(i)]|^2\right) \\ &= E(\overline{\mu}(i)|e_a(i)|^2) - \underbrace{2\mu E\left[e_a\left(R\,\text{sign}\,y(i) - y(i)\right)\right]\overline{\mu}(i)}_{H'} \\ &\quad \underbrace{+\mu^2 E\left[\left(R\,\text{sign}\,y(i) - y(i)\right)^2\right]\overline{\mu}(i)}_{I'}. \end{aligned}$$

We consider the case where $h_D = [0, \cdots, 0, 1, 0, \cdots, 0]$. Similar results can be obtained when $h_D = [0, \cdots, 0, -1, 0, \cdots, 0]$. Then, $e_a = s - y$. Let us first examine the term H' by replacing y with $s - e_a$. With Assumption I.2 (where $\mu^2||u||^2$ is independent of y and e_a), the term H' can be written as

$$H' = \frac{2\mu}{E||u||^2}\left[REe_a \,\text{sign}\,(y) - Ee_a s + Ee_a^2\right].$$

With Assumptions I.3 and I.4, we get $Ee_a \,\text{sign}\,(y) = Ee_a E \,\text{sign}\,(y) = 0$. In addition, from Assumption I.1, and because $Es = 0$, we get $Ee_a s = 0$. Hence, the term H' can be simplified to

$$H' = 2\frac{\mu}{E||u||^2}E|e_a|^2.$$

Now, we evaluate the term I'. From Assumption I.2, I' can be written as

$$I' = \frac{\mu^2}{E||u||^2}\,E(R\,\text{sign}\,(y) - s + e_a)^2.$$

Expanding this expression, we get

$$\begin{aligned} I' = \frac{\mu^2}{E||u||^2}\,E\,&\big(R^2 + s^2 + e_a^2 - 2Rs\,\text{sign}\,(y) - 2se_a \\ &+2Re_a\,\text{sign}\,(y)\big)\,. \end{aligned}$$

For small enough e_a, we can write $\text{sign}\,(y) = \text{sign}\,(s)$ so that $Es \cdot \,\text{sign}\,(y) = E|s|$, and

$$I' = \frac{\mu^2}{E||u||^2}\,E\left(R^2 + s^2 + e_a^2 - 2R|s|\right).$$

µAgain, when and e_a^2 are sufficiently small, the term $\mu^2/(E||u||^2) \cdot Ee_a^2$ can be ignored, and hence

$$I' \approx \frac{\mu^2}{E||u||^2}\,E(R^2 + s^2 - 2R|s|).$$

Then, $H' = I'$ leads to (37). This expression for the MSE is non-negative for any R because

$$R^2 + E|s|^2 - 2RE|s| = E(R - |s|)^2 \geq 0.$$

Equality occurs only for constant modulus signals. □

REFERENCES

[1] R. Liu and L. Tong, Eds., "Special issue on blind identification and equalization," *Proc. IEEE*, vol. 86, Oct. 1998.

[2] J. R. Treichler and I. Fijalkow, "Fractionally spaced equalizers: How long should they be?," *IEEE Signal Processing Mag.*, vol. 13, pp. 65–81, May 1996.

[3] C. R. Johnson *et al.*, "Blind equalization using the constant modulus criterion: A review," *Proc. IEEE*, vol. 86, pp. 1927–1950, Oct. 1998.

[4] A. V. Keerthi, A. Mathur, and J. Shynk, "Misadjustment and tracking analysis of the constant modulus array," *IEEE Trans. Signal Processing*, vol. 46, pp. 51–58, Jan. 1998.

[5] O. W. Kwon, C. K. Un, and J. C. Lee, "Performance of constant modulus adaptive digital filters for interference cancellation," *Signal Process.*, vol. 26, no. 2, pp. 185–196, Feb. 1992.

[6] N. R. Yousef and A. H. Sayed, "A unified approach to the steady-state and tracking analyzes of adaptive filtering algorithms," in *Proc. 4th IEEE-EURASIP Int.Workshop Nonlinear Signal Image Process. (NSIP)*, vol. 2, Antalya, Turkey, June 1999, pp. 699–703.

[7] ——, "A feedback analysis of the tracking performance of blind adaptive equalization algorithms," in *Proc. Conf. Decision Contr.*, Phoenix, AZ, Dec. 1999.

[8] ——, "Tracking analysis of the LMF and LMMN adaptive algorithms," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Oct. 1999.

[9] C. K. Chan and J. J. Shynk, "Stationary points of the constant modulus algorithm for real Gaussian signals," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 38, pp. 2176–2181, Dec. 1990.

[10] J. J. Shynk, R. P. Gooch, G. Krisnamurthy, and C. K. Chan, "A comparative performance study of several blind equalization algorithms," in *Proc. SPIE Conf. Adv. Signal Process.*, vol. 1565, 1991, pp. 102–117.

[11] N. J. Bershad and S. Roy, "Performance of the 2-2 constant modulus (CM) adaptive algorithm for Rayleigh fading sinusoids in Gaussian noise," in *Proc. ICASSP*, Albuquerque, NM, Apr. 1990, pp. 1675–1678.

[12] H. H. Zeng and L. Tong, "The MSE performance of constant modulus receivers," in *Proc. ICASSP*, Munich, Germany, Apr. 1997, pp. 3577–3580.

[13] I. Fijakow, C. E. Manlove, and C. R. Johnson, "Adaptive fractionally spaced blind CMA equalization: Excess MSE," *IEEE Trans. Signal Processing*, vol. 46, pp. 227–231, Jan. 1998.

[14] G. Ungerboeck, "Fractional tap-spacing equalizer and consequences for clock recovery in data modems," *IEEE Trans. Commun.*, vol. –COMM24, pp. 856–864, Aug. 1976.

[15] I. Fijalkow and F. L. de Victoria, "Adaptive fractionally spaced blind equalization," in *Proc. IEEE Signal Process. Workshop*, Yosemite, CA, Oct. 1994, pp. 257–260.

[16] Y. Li and Z. Ding, "Global convergence of fractionally spaced Godard (CMA) adaptive equalizers," *IEEE Trans. Signal Processing*, vol. 44, pp. 818–826, Apr. 1996.

[17] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. 28, no. 11, pp. 1867–1875, Nov. 1980.

[18] J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. –ASSP31, pp. 349–472, Apr. 1983.

[19] Z. Ding, C. R. Johnson, and R. A. Kennedy, "On the (non)existence of undesirable equilibria of Godard blind equalizers," *IEEE Trans. Signal Processing*, vol. 40, pp. 2425–2432, Oct. 1992.

[20] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[21] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.

[22] O. Macchi, *Adaptive Processing: The LMS Approach with Applications in Transmission*. New York: Wiley, 1995.

[23] O. Macchi and E. Eweda, "Second-order convergence analysis of stochastic adaptive linear filtering," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 76–85, Jan. 1983.

[24] J. K. Kim and L. D. Davisson, "Adaptive linear estimation for stationary M-dependent processes," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 23–31, Jan. 1975.

[25] J. E. Mazo, "On the independence theory of equalizer convergence," *Bell Syst. Tech. J.*, vol. 58, pp. 963–993, May/June 1979.

[26] A. H. Sayed and M. Rupp, "A time domain feedback analysis of adaptive gradient algorithms via the small gain theorem," in *Proc. SPIE*, vol. 2563, San Diego, CA, July 1995, pp. 458–469.

[27] M. Rupp and A. H. Sayed, "A time-domain feedback analysis of filtered-error adaptive gradient algorithms," *IEEE Trans. Signal Processing*, vol. 44, pp. 1428–1439, June 1996.

[28] A. H. Sayed and M. Rupp, "An l_2-stable feedback structure for nonlinear adaptive filtering and identification," *Automatica*, vol. 33, no. 1, pp. 13–30, Jan. 1997.

[29] ——, "Robustness issues in adaptive filtering," in *DSP Handbook*. Boca Raton, FL: CRC, 1998, ch. 20.

[30] V. Solo and X. Kong, *Adaptive Signal Processing Algorithms*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[31] H. J. Kushner, *Approximation and Weak Convergence Methods for Random Processes with Applications to Stochastic System Theory*. Cambridge, MA: MIT Press, 1984.

[32] A. Benveniste, M. Metivier, and P. Priouret, *Adaptive Algorithms and Stochastic Approximations*. New York: Springer-Verlag, 1990.

[33] E. Eweda, "Comparison of RLS, LMS, and sign algorithms for tracking randomly time-varying channels," *IEEE Trans. Signal Processing*, vol. 42, pp. 2937–2944, Nov. 1994.

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

[35] K. Hilal and P. Duhamel, "A convergence study of the constant modulus algorithm leading to a normalized-CMA and a block-normalized-CMA," in *Proc. EUSIPCO*, Brussels, Belgium, Aug. 1992, pp. 135–138.

[36] C. B. Papadias and D. T. M. Slock, "On the convergence of normalized constant modulus algorithms for blind equalization," in *Proc. Int. Conf. Digital Signal Processing*, Nicosia, Cyprus, July 1993, pp. 245–250.

[37] ——, "Normalized sliding window constant modulus and decision-directed algorithms: A link between blind equalization and classical adaptive filtering," *IEEE Trans. Signal Processing*, vol. 45, pp. 231–235, Jan. 1997.

Junyu Mai received the B.S. degree in biotechnology from Tsinghua University, Tsinghua, Taiwan, R.O.C., in 1995 and the M.S. degrees in biochemistry and in electrical engineering from the University of California, Los Angeles (UCLA), in 1996 and 1998, respectively.

She is currently a Senior Research Scientist with the Advanced Research Department, St. Jude Medical Cardiac Rhythm Management Division, Sylmar, CA, where she is working on techniques for congestive heart failure monitoring, heart sound detection and analysis, and body position detection. Her research interests are in adaptive algorithms and biomedical applications.

–Ms. Mai was the recipient of the 19981999 Outstanding M.S. Student Award from the Electrical Engineering Department at UCLA.

Ali H. Sayed (SM'99) received the Ph.D. degree in electrical engineering in 1992 from Stanford University, Stanford, CA.

He is Associate Professor of electrical engineering at the University of California, Los Angeles. He has more than 120 journal and conference publications, is a coauthor of *Indefinite Quadratic Estimation and Control* (Philadelphia, PA: SIAM, 1999) and *Linear Estimation* (Englewood Cliffs, NJ: Prentice-Hall, in press), and a coeditor of *Fast Reliable Algorithms for Matrices with Structure* (Philadelphia, PA: SIAM, 1999). He has contributed several articles to engineering and mathematical encyclopedias and handbooks. His research interests span several areas including adaptive and statistical signal processing, linear and nonlinear filtering and estimation, linear and nonlinear system and control theories, interplays between signal processing and control methodologies, and reliable and efficient algorithms for large-scale structured computations.

Dr. Sayed is a member of the editorial board of the *SIAM Journal on Matrix Analysis and Its Applications*, has served as coeditor of special issues of the journal *Linear Algebra and Its Applications*, and is past Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He has served on the program committees of several international meetings. He is also a recipient of the 1996 IEEE Donald G. Fink Award.