

Blind Equalization Using the Constant Modulus Criterion: A Review

C. RICHARD JOHNSON, JR., FELLOW, IEEE, PHILIP SCHNITER, THOMAS J. ENDRES, MEMBER, IEEE, JAMES D. BEHM, DONALD R. BROWN, AND RAÚL A. CASAS

Invited Paper

This paper provides a tutorial introduction to the constant modulus (CM) criterion for blind fractionally spaced equalizer (FSE) design via a (stochastic) gradient descent algorithm such as the constant modulus algorithm (CMA). The topical divisions utilized in this tutorial can be used to help catalog the emerging literature on the CM criterion and on the behavior of (stochastic) gradient descent algorithms used to minimize it.

Keywords— Adaptive equalizers, blind deconvolution, blind equalization, constant modulus algorithm (CMA), digital communication, equalizers, intersymbol interference, least mean square methods.

I. INTRODUCTION

Information-bearing signals transmitted between remote locations often encounter a signal-altering physical channel. Examples of common physical channels include coaxial, fiber optic, or twisted-pair cable in wired communications and the atmosphere or ocean in wireless communications. Each of these physical channels may cause signal distortion, including echoes and frequency-selective filtering of the transmitted signal. In digital communications, a critical manifestation of distortion is intersymbol interference (ISI), whereby symbols transmitted before and after a given symbol corrupt the detection of that symbol. All physical channels (at high enough data rates) tend to exhibit ISI. The presence of ISI is readily observable in the sampled impulse response of a channel; an impulse response corresponding to a lack of ISI contains a single spike of width less than the time between symbols. An example of a terrestrial microwave channel impulse response [obtained from the Rice

Manuscript received May 13, 1997; revised February 27, 1998. This work was supported in part by National Science Foundation Grant MIP-9509011 and Applied Signal Technology.

C. R. Johnson, Jr., P. Schniter, D. R. Brown, and R. A. Casas are with the School of Electrical Engineering, Cornell University, Ithaca, NY 14853 USA (e-mail: johnson@ee.cornell.edu; schniter@ee.cornell.edu; brownr@ee.cornell.edu; raulc@ee.cornell.edu).

T. J. Endres is with Sarnoff Digital Communications, Newtown, PA 18940 USA (e-mail: endres@erols.com).

J. D. Behm is with the Department of Defense, Ft. Meade, MD 20855 USA (e-mail: jimbehm@toad.net).

Publisher Item Identifier S 0018-9219(98)06972-2.

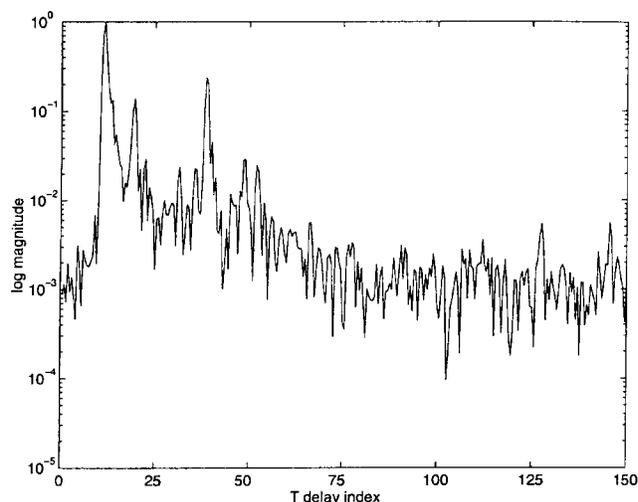


Fig. 1. Terrestrial microwave channel impulse response magnitude, $1/T = 30 \times 10^6$ symbols/s (SPIB channel #3).

University Signal Processing Information Base (SPIB)¹] is shown in Fig. 1.

Linear channel equalization, an approach commonly used to counter the effects of linear channel distortion, can be viewed as the application of a linear filter (i.e., the equalizer) to the received signal. The equalizer attempts to extract the transmitted symbol sequence by counteracting the effects of ISI, thereby improving the probability of correct symbol detection.

Since it is common for the channel characteristics to be unknown (e.g., at startup) or to change over time, the preferred embodiment of the equalizer is a structure adaptive in nature. Classical equalization techniques employ a time-slot (recurring periodically for time-varying situations) during which a training signal, known in advance by the receiver, is transmitted. The receiver adapts the equalizer (e.g., via LMS [6], [27]) so that its output closely matches the known

¹This microwave channel database resides at <http://spib.rice.edu/spib/microwave.html>.

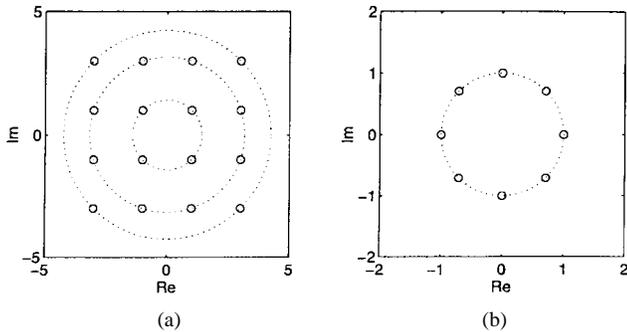


Fig. 2. (a) Nonconstant modulus source constellation (16-QAM) versus (b) CM source constellation (8-PSK).

reference training signal. Since the inclusion of such signals sacrifices valuable channel capacity, adaptation without resort to training, i.e., blind adaptation, is preferred. The most studied and implemented blind adaptation algorithm of the 1990's is the constant modulus algorithm (CMA).

CMA seeks to minimize a cost defined by the CM criterion. The CM criterion penalizes deviations in the modulus (i.e., magnitude) of the equalized signal away from a fixed value. In certain ideal conditions, minimizing the CM cost can be shown to result in perfect (zero-forcing) equalization of the received signal. Remarkably, the CM criterion can successfully equalize signals characterized by source alphabets not possessing a constant modulus [e.g., 16-quadrature amplitude modulation (QAM)], as well as those possessing a constant modulus (e.g., 8-PSK) (see Fig. 2). This paper attempts to explore the behavior of CMA by considering the similarities between the CM and mean-squared error (MSE) criteria. This relationship is important because of well-known connections between MSE and the actual quantity we desire to minimize: probability of bit error (e.g., see the discussion in [5]).

Plotting the CM cost versus the equalizer coefficients results in a surface referred to as the CM cost surface. Stochastic gradient descent (SGD) algorithms [6], [13] attempt to minimize the CM cost by starting at some location on the surface and following the trajectory of steepest descent. The CM cost surface characteristics are important because they can be used to understand the behavior of any SGD algorithms that attempt to minimize the CM cost, such as CMA. Specifically, these characteristics lend insight into the channel, equalizer, and source properties which affect SGD behavior.

The success of a stochastic gradient descent equalizer adaptation algorithm is dependent on a certain amount of stationarity in the received process. Thus, throughout the paper, we restrict our focus to stationary source and noise processes and to channels whose impulse response is fixed or slowly² time varying.

A. History

In the literature, blind equalization algorithms blossomed in the 1980's. The two principal precursors are Lucky's

²Here "slow" is considered relative to the tracking speed of the SGD algorithm.

blind decision-direction algorithm [11] and Sato's algorithm [19]. What we term the CM criterion was introduced for blind equalization of QAM signals in [29] and of pulse-amplitude modulation (PAM) and FM signals in [30]. By the end of the 1980's blind equalizers were commercialized for microwave radio [9]. By the mid 1990's blind equalizers were realized in very large scale integration (VLSI) for high definition television (HDTV) set-top cable demodulators [23]. The current explosion of interest in the CM criterion stems from blind processing applications in emerging wireless communication technology (e.g., blind equalization, blind source separation, and blind antenna steering) and from CMA's record of practical success.

B. Our Mission

This paper is intended to be a resource to both readers experienced in blind equalization as well as those new to the subject. In a tutorial style, Section I-C provides background in fractionally spaced equalizer (FSE) modeling and design. (For baud-spaced equalizer (BSE) design, we refer the interested reader to a variety of classical references, e.g., [5], [6], [10], and [16].) Section II then illustrates several low-dimensional examples that help to characterize the behavior of FSE's adapted under the constant modulus criterion.

In Section III we construct a categorization of literature focusing on the application of the CM criterion to blind equalization. The annotated bibliography in Appendix III catalogs the existing literature according to the classifications of Section III, providing the reader with a valuable tool for further research. Our attempt to be exhaustive is justified only by the relative infancy of the subfield; evidence of the emerging status of this literature is seen in the wealth of conference papers in the bibliography of Appendix III.

Following the introductory FSE tutorial, Section I-E presents a novel view of classical nonblind adaptive equalization that illuminates the connection between the MSE and CM criteria. Specifically, the LMS-with-training strategy requires preselection of a design variable, namely training sequence delay, that may lead to a potentially suboptimal solution. The delay-optimized MSE (a function of equalizer parameters only) yields a cost surface (see Fig. 7) for which a simple LMS-like parameter update algorithm is not known to exist. Remarkably, the CM criterion offers a proxy for this surface for which there exists a (blind) parameter update algorithm, namely, CMA.

C. Fractionally Spaced Linear Equalization

In this section we describe the fractionally spaced equalization scenario and present some fundamental results regarding minimum MSE (i.e., Wiener [6]) equalizers. This material is primarily intended to provide background and context. For simplicity, our focus is restricted to a $T/2$ -spaced FSE, where T denotes the baud, or symbol, duration. All results are extendible to the more general T/N -spaced case. Examples of seminal work on fractionally spaced

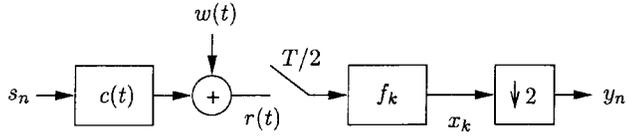


Fig. 3. Baseband model of single-channel communication system with $T/2$ -spaced receiver.

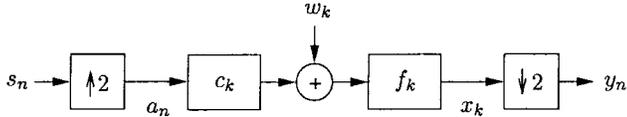


Fig. 4. Multirate system model.

equalization include [4], [12], [14], and [26], while more comprehensive references are [5] and [18].

1) *Multirate and Multichannel System Models:* Consider the single-channel model illustrated in Fig. 3. A (possibly complex-valued) T -spaced symbol sequence $\{s_n\}$ is transmitted through a pulse shaping filter, modulated onto a propagation channel, and demodulated. We assume all processing between the transmitter and receiver is linear and time invariant (LTI) and can thus be described by the continuous-time impulse response $c(t)$. The received signal $r(t)$ is also corrupted by additive channel noise, whose baseband equivalent we denote by $w(t)$. The received signal is then sampled at $T/2$ -spaced intervals and filtered by a $T/2$ -spaced finite impulse response (FIR) equalizer of length $2N$. (An even length is chosen for notational simplicity.) This filtering can be regarded as a convolution of the sampled received sequence with the equalizer coefficients f_k . Finally, the FSE output $\{x_k\}$ is decimated by a factor of two to create the T -spaced output sequence $\{y_n\}$. Decimation is accomplished by disregarding alternate samples, thus producing the baud-spaced “soft decisions” y_n . We note that, in general, all quantities are complex valued. For clarity, we reserve the index n for baud-spaced quantities and the index k for fractionally spaced quantities throughout the paper.

Appendix I derives the equivalence between the continuous-time model in Fig. 3 and the discrete time models in Figs. 4 and 5, both constructed using $T/2$ -spaced samples of $c(t)$ and $w(t)$. Fig. 4 depicts the multirate model while Fig. 5 depicts the multichannel model. Though our derivation of the discrete-time models is based on the single-channel system in Fig. 3, the equivalence between the multirate and multichannel models suggests that we could have based our model on a two-sensor (T -sampled) communication system instead. For a concise discussion on the equivalence between temporal and spatial diversity, see [15].

The multirate model of Fig. 4 uses the discrete-time fractionally spaced channel coefficients $c_k = c(k(T/2))$ and the discrete-time random process $w_k = w(k(T/2))$. The multichannel model of Fig. 5 subdivides these sample sequences into even and odd baud-spaced counterparts (of relative delay $T/2$), so that $c_n^{\text{even}} = c_{2n}$ and $c_n^{\text{odd}} =$

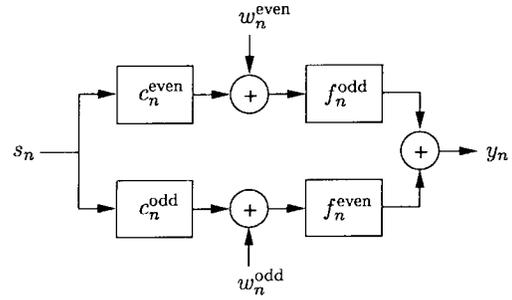


Fig. 5. Multichannel system model.

c_{2n+1} for $n = 0, 1, 2, \dots$. In a similar manner, the FSE coefficients are partitioned as $f^{\text{even}} = f_{2n}$ and $f^{\text{odd}} = f_{2n+1}$.

Given a fractionally spaced channel of finite³ length $2M$, we can collect the even and odd sets of equalizer and channel coefficients into column vectors

$$\begin{aligned} \mathbf{f}_e &= [f_0, f_2, f_4, \dots, f_{2N-2}]^t \\ &= [f_0^{\text{even}}, f_1^{\text{even}}, f_2^{\text{even}}, \dots, f_{N-1}^{\text{even}}]^t \\ \mathbf{f}_o &= [f_1, f_3, f_5, \dots, f_{2N-1}]^t \\ &= [f_0^{\text{odd}}, f_1^{\text{odd}}, f_2^{\text{odd}}, \dots, f_{N-1}^{\text{odd}}]^t \\ \mathbf{c}_e &= [c_0, c_2, c_4, \dots, c_{2M-2}]^t \\ &= [c_0^{\text{even}}, c_1^{\text{even}}, c_2^{\text{even}}, \dots, c_{M-1}^{\text{even}}]^t \\ \mathbf{c}_o &= [c_1, c_3, c_5, \dots, c_{2M-1}]^t \\ &= [c_0^{\text{odd}}, c_1^{\text{odd}}, c_2^{\text{odd}}, \dots, c_{M-1}^{\text{odd}}]^t. \end{aligned} \quad (1)$$

It is possible to form the (baud-spaced) impulse response of the linear system relating s_n to y_n using a pair of $P \times N$ baud-spaced convolution matrices \mathbf{C}_e and \mathbf{C}_o , where $P = M + N - 1$.

$$\mathbf{C}_e = \begin{bmatrix} c_0^{\text{even}} & & & & \\ c_1^{\text{even}} & c_0^{\text{even}} & & & \\ \vdots & c_1^{\text{even}} & & & \\ c_{M-1}^{\text{even}} & \vdots & \ddots & c_0^{\text{even}} & \\ & c_{M-1}^{\text{even}} & & c_1^{\text{even}} & \\ & & & \vdots & \\ & & & & c_{M-1}^{\text{even}} \end{bmatrix}$$

$$\mathbf{C}_o = \begin{bmatrix} c_0^{\text{odd}} & & & & \\ c_1^{\text{odd}} & c_0^{\text{odd}} & & & \\ \vdots & c_1^{\text{odd}} & & & \\ c_{M-1}^{\text{odd}} & \vdots & \ddots & c_0^{\text{odd}} & \\ & c_{M-1}^{\text{odd}} & & c_1^{\text{odd}} & \\ & & & \vdots & \\ & & & & c_{M-1}^{\text{odd}} \end{bmatrix}. \quad (2)$$

³In practice, we would consider the fractionally spaced channel to be of “finite length” M if the response magnitude can be said to decay below some sufficiently small threshold for all time $t \geq M(T/2)$.

for a particular choice of delay (δ). We will see that this criterion can be interpreted as the best compromise between ISI and noise amplification in a minimum mean-squared error (MMSE) sense.

To formulate this error criterion more precisely, we collect the P previous T -spaced elements of the source sequence into the vector

$$\mathbf{s}(n) = [s_n, s_{n-1}, s_{n-2}, \dots, s_{n-(P-1)}]^t \quad (11)$$

and the last $2N$ fractionally sampled values of noise into vector $\mathbf{w}(n)$

$$\mathbf{w}(n) = [w_{n-1}, w_{n-3}, w_{n-5}, \dots, w_{n-(2N-1)}, w_n, w_{n-2}, w_{n-4}, \dots, w_{n-(2N-2)}]^t \quad (12)$$

where the collection of even noise samples follows the collection of odd noise samples to be consistent with our definitions of \mathbf{C} and \mathbf{f} in (3). [Note, however, that this particular ordering of samples in the noise vector is inconsequential when assuming an independent identically distributed (i.i.d.) noise process.] With these quantities, the n th equalizer output $y_n = y(nT + (T/2))$ can be written compactly as

$$y_n = \mathbf{s}^t(n)\mathbf{C}\mathbf{f} + \mathbf{w}^t(n)\mathbf{f} \quad (13)$$

yielding an expression for the recovery error

$$e_n = \mathbf{s}^t(n)(\mathbf{C}\mathbf{f} - \mathbf{h}_\delta) + \mathbf{w}^t(n)\mathbf{f}. \quad (14)$$

Under the assumption that the noise and source processes are i.i.d. and jointly uncorrelated, with respective variances σ_w^2 and σ_s^2 , the expected value of the magnitude-squared recovery error becomes

$$E\{|e_n|^2\} = (\mathbf{C}\mathbf{f} - \mathbf{h}_\delta)^H(\mathbf{C}\mathbf{f} - \mathbf{h}_\delta)\sigma_s^2 + \mathbf{f}^H\mathbf{f}\sigma_w^2. \quad (15)$$

(Appendix I-D discusses the independence assumption regarding fractionally sampled channel noise.) Note that (15) is proportional to the source-power-normalized MSE cost function

$$J_{\text{MSE}} = (\mathbf{C}\mathbf{f} - \mathbf{h}_\delta)^H(\mathbf{C}\mathbf{f} - \mathbf{h}_\delta) + \lambda\mathbf{f}^H\mathbf{f} \quad (16)$$

where $\lambda = \sigma_w^2/\sigma_s^2$. In terms of $\mathbf{A} = \mathbf{C}^H\mathbf{C} + \lambda\mathbf{I}$, the technique of “completing the square” yields

$$J_{\text{MSE}} = (\mathbf{f} - \mathbf{A}^{-1}\mathbf{C}^H\mathbf{h}_\delta)^H\mathbf{A}(\mathbf{f} - \mathbf{A}^{-1}\mathbf{C}^H\mathbf{h}_\delta) - \mathbf{h}_\delta^H\mathbf{C}\mathbf{A}^{-1}\mathbf{C}^H\mathbf{h}_\delta + \mathbf{h}_\delta^H\mathbf{h}_\delta. \quad (17)$$

Note that \mathbf{A} is positive definite for $\lambda > 0$.

Equation (17) indicates that the equalizer parameter vector minimizing J_{MSE} is

$$\mathbf{f}^\dagger = \mathbf{A}^{-1}\mathbf{C}^H\mathbf{h}_\delta \quad (18)$$

and it follows that the \mathbf{f} -optimal MSE

$$\min_{\mathbf{f}} J_{\text{MSE}}(\mathbf{f}, \delta) = J_{\text{MSE}}(\mathbf{f}^\dagger, \delta) \quad (19)$$

$$= \mathbf{h}_\delta^H(\mathbf{I} - \mathbf{C}\mathbf{A}^{-1}\mathbf{C}^H)\mathbf{h}_\delta \quad (20)$$

remains a function of system delay δ . We make this property explicit by adopting the notation $J_{\text{MSE}}(\mathbf{f}, \delta)$. It follows from (8) and (20) that the optimum delay δ^\dagger

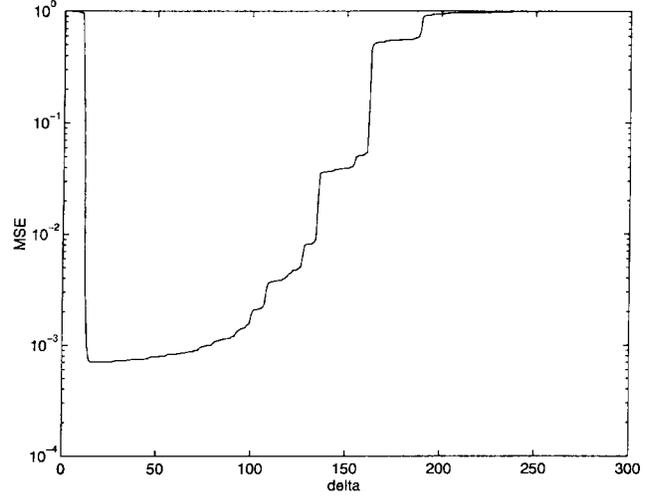


Fig. 6. \mathbf{f} -optimal MSE, $J_{\text{MSE}}(\mathbf{f}^\dagger, \delta)$, versus T -spaced delay δ for the channel of Fig. 1 and 30 dB SNR using 300-tap FSE.

corresponds to the index of the minimum diagonal element of $\mathbf{I} - \mathbf{C}\mathbf{A}^{-1}\mathbf{C}^H$ [7]. This is written formally as

$$\delta^\dagger = \arg \min_{\delta} \left\{ \left[\mathbf{I} - \mathbf{C}(\mathbf{C}^H\mathbf{C} + \lambda\mathbf{I})^{-1}\mathbf{C}^H \right]_{\delta, \delta} \right\}. \quad (21)$$

For a $T/2$ -spaced FSE with 300 taps and an SNR ($= 1/\lambda$) of 30 dB, Fig. 6 plots $J_{\text{MSE}}(\mathbf{f}^\dagger, \delta)$ versus δ for the “typical” impulse response of Fig. 1. Note the degree to which δ can affect MSE performance.

We conclude that proper preselection of δ is important for equalizer-based minimization of $J_{\text{MSE}}(\mathbf{f}, \delta)$. This idea of fixed- δ optimization is of particular relevance because it describes the typical adaptive equalization scenario when a training signal is available [17].

E. An Amalgamated MSE Cost Function

When the source is differentially encoded [5], knowledge of absolute phase is not required for symbol detection. For example, either $y_n = s_{n-\delta}$ or $y_n = -s_{n-\delta}$ (for all n) would form an acceptable output sequence for differentially encoded binary phase-shift keying (BPSK). (For complex-valued source alphabets such as QAM, we allow $y_n = e^{j(\pi/2)m}s_{n-\delta}$ for fixed $m \in \{0, 1, 2, 3\}$.) Therefore, an acceptable system impulse response can include a fixed phase shift in addition to a bulk system delay δ . With this in mind, we construct a phase- and delay-optimized amalgamated cost function $J_A(\mathbf{f})$

$$J_A(\mathbf{f}) = \min_{\delta, \rho} \{ (\mathbf{C}\mathbf{f} - \rho\mathbf{h}_\delta)^H(\mathbf{C}\mathbf{f} - \rho\mathbf{h}_\delta) + \lambda\mathbf{f}^H\mathbf{f} \} \quad (22)$$

where ρ is one of the set of allowable phase shifts (e.g., $\{+1, -1\}$ for real-valued PAM).

J_A is a multimodal fabrication, bearing similarity to a $(2N + 1)$ -dimensional egg carton. A surface plot appears in Fig. 7 for well-behaved $T/2$ -spaced channel \mathbf{c}_1 defined in Table 1. By “well-behaved” we mean that \mathbf{c}_1 has no common or nearly common subchannel roots. Fig. 7 indicates that if we minimize $J_A(\mathbf{f})$ by a gradient descent strategy, then the initial value of \mathbf{f} will determine the values

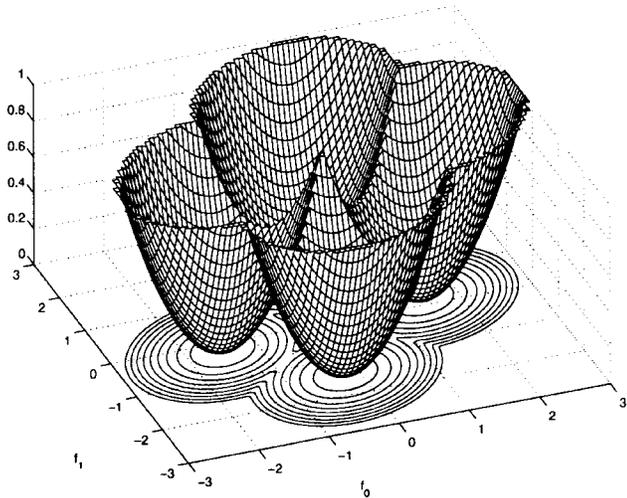


Fig. 7. J_A for well-behaved channel c_1 and no noise in equalizer (\mathbf{f}) space.

Table 1 Summary of Channels Used for Two-Tap FSE Examples

Name	$T/2$ -spaced Impulse Response	Classification
c_1	[-0.0901, 0.6853, 0.7170, -0.0901]	well-behaved ⁷
c_2	[1.0, -0.5, 0.2, 0.3]	well-behaved
c_3	[-0.0086, 0.0101, 0.9999, -0.0086]	nearly-common subchannel roots
c_4	[1.0, -0.5, 0.2, 0.3, -0.2, -0.15]	undermodelled

of δ and ρ to which the descent scheme will asymptotically converge. In other words, optimization of $J_A(\mathbf{f})$ by gradient descent accomplishes preselection of δ via choice of \mathbf{f} -initialization.

Section II attests to the claim that *the CM criterion serves as a close proxy to J_A , which is robust under typical operating conditions*. For a preview, compare the CM cost surface in Fig. 8 to the amalgamated MSE surface in Fig. 7 for the same channel c_1 . As such, the CM criterion offers a performance metric that bears many similarities to MSE but which is capable of minimization by (stochastic) gradient descent schemes conducted blindly with respect to the transmitted symbols.

With our tutorial orientation, Section II restricts focus to a two-tap FSE design task that permits visualization of equalizer-parameter-space cost-contour plots illustrating various properties of the CM cost function J_{CM} . In particular, we can isolate an “ideal zero-cost” situation where the stationary points in J_{CM} and J_A match exactly and where the minima achieve zero cost. This special case requires several assumptions not often satisfied in practice. We will examine examples of CM-adapted FSE behavior conducted under violations of these requirements for ideal zero-cost equalization. This implicit taxonomy will be used in Section III to provide an overview of the literature citations in the annotated bibliography of Appendix III.

II. TWO-TAP ILLUSTRATIVE EXAMPLES

The shape of the cost surface defining a particular stochastic gradient algorithm often lends great insight into the

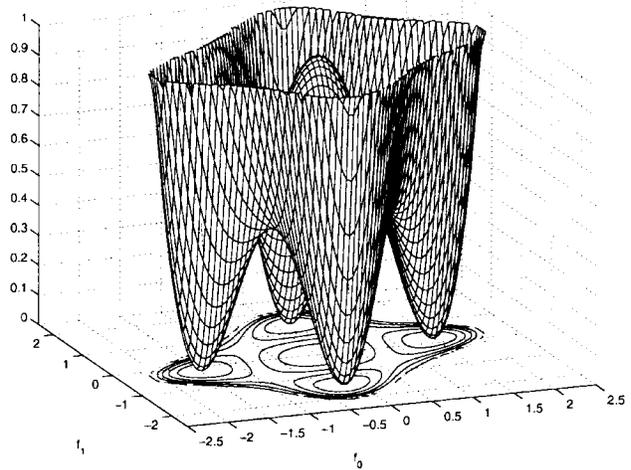


Fig. 8. J_{CM} for well-behaved channel c_1 and no noise in equalizer (\mathbf{f}) space.

expected behavior of that algorithm. With this in mind, we embark on a tutorial study of the cost surface defined by the CM criterion and descended by CMA. First, however, consider the following list of features characterizing a generic (stochastic) gradient descent algorithm.

- Far from a stationary point, the gradient (i.e., first derivative) of the cost surface determines local convergence rate.
- Near a stationary point, the local curvature (i.e., second derivative) of the cost surface determines local convergence rate.
- Local minima with nonzero cost induce excess steady-state error in stochastic gradient descent algorithms with nonvanishing step-sizes.
- Multimodal surfaces may exhibit local minima of varying cost, thus linking initialization to achievable asymptotic performance.
- “Poor” initialization on a multimodal surface can lead a trajectory into temporary capture by (one or more) saddle points, resulting in arbitrarily slow convergence to a minimum.
- Nontrivial deformations of a multimodal surface relocate each saddle point and alter the region of attraction associated with each local minima.

The following sections combine low-dimensional examples with the well-known characteristics above to formulate an intuitive understanding of the CM criterion and its connection to the MSE criterion.

A. Two-Tap Equalizer Design Equations

As discussed in Section I-C, satisfaction of the “length and zero” conditions ensures an exact solution to the zero-forcing equation $\mathbf{h}_s = \mathbf{C}\mathbf{f}$. For a two-tap $T/2$ -spaced FSE, the length condition is satisfied for channels with impulse responses $[c_0, c_1, c_2, c_3]$ and shorter. For a length-four channel, the root condition is satisfied when the even and odd subchannel polynomials $C_{\text{even}}(z^{-1}) = c_0 + c_2z^{-1}$ and $C_{\text{odd}}(z^{-1}) = c_1 + c_3z^{-1}$ have distinct roots.

In this case, (3) specifies that the FSE design quantities take the following form:

$$\mathbf{C} = \begin{bmatrix} c_1 & c_0 \\ c_3 & c_2 \end{bmatrix} \quad \mathbf{f} = \begin{bmatrix} f_0 \\ f_1 \end{bmatrix}. \quad (23)$$

Since \mathbf{h}_δ has one nonzero coefficient, the zero-forcing equalizer will be proportional to either the first or the second column of \mathbf{C}^{-1} . Thus, all four channel parameters enter into the design of \mathbf{f} ; the sub-equalizers of Fig. 5 are not simply inverses of their respective subchannels.

B. Introduction to the CM Cost Function

The CM cost function can be motivated using the temporary assumption that the source is binary valued (± 1). In this case, s_n has a constant squared-modulus of one ($|s_n|^2 = 1$). Under perfect symbol recovery we know that the output y_n has the same CM property and can thus imagine a cost that penalizes deviations from this output condition. This, in fact, defines the CM cost function for a BPSK source

$$J_{\text{CM}|_{\text{BPSK}}} = E\left\{(1 - |y_n|^2)^2\right\}.$$

Appendix II presents more general versions of the CM cost function and derives expressions for J_{CM} in terms of channel parameters, particular source and noise statistics, and equalizer coefficients.

The leap of faith, first espoused by [29], is the application of J_{CM} to a multilevel (i.e., nonconstant modulus) source. Reference [29], which addressed baud-spaced blind equalization via minimization of J_{CM} , makes the first observation concerning the proximity of the J_{CM} and J_{A} minima.

It should also be noted that the equalizer coefficients minimizing the dispersion functions closely approximate those which minimize the mean squared error.

This is remarkable because an approximation of J_{CM} can be formed solely from the equalizer output y_n ; no training signal is required to compose an accurate gradient approximation for use in a stochastic gradient minimization algorithm such as CMA [30]. It is worth noting that the phase-independent nature of J_{CM} has its own advantages in modem design [24].

C. Illustrative Cost Surface Examples

The following sections present mesh and contour plots of the CM cost surface for a two-tap $T/2$ -spaced FSE under various operating conditions. Refer to Table 1 for definitions of the various channels used in our experiments. In all contour plots, the asterisks (*) indicate the locations of global MSE (i.e., J_{A}) minima while the crosses (x) indicate the locations of local MSE minima. Recall that different pairs of MSE minima (reflected through the origin) correspond to different values of system delay, while the two elements composing each pair correspond to the two

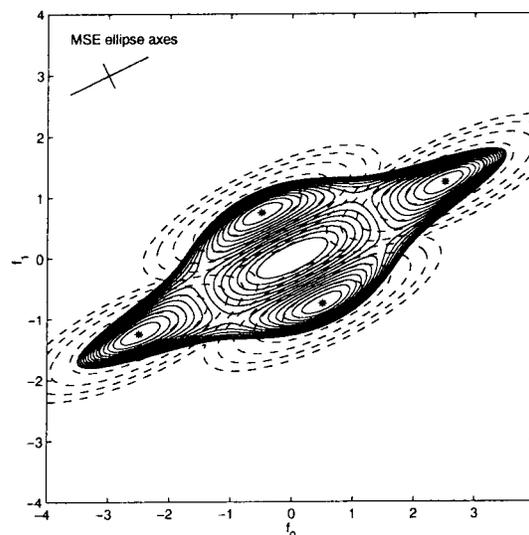


Fig. 9. J_{CM} contours (solid) for well-behaved channel c_2 and no noise, with J_{A} overlay (dashed) and global MSE minima marked by “*” in equalizer (\mathbf{f}) space.

choices of system polarity.⁶ Thus, the asterisks mark the MMSE equalizers of optimum system delay. The “MSE ellipse axes” appearing in the upper left corner of each contour plot indicate the orientation and eccentricity of the elliptical MSE contours (see Fig. 9).

All quantities in the experiments are real-valued. Unless otherwise noted, the source used was zero-mean and i.i.d. with alphabet $\{-1, 1\}$.

1) *Ideal Zero-Cost Equalization:* For a well-behaved⁷ channel c_1 in the absence of channel noise, Fig. 8 plots J_{CM} in equalizer space. Recall that Fig. 7 plots J_{A} for the same noiseless channel. For a different well-behaved and noiseless channel c_2 , Fig. 9 superimposes the corresponding J_{CM} and J_{A} cost contours. Note the symmetry (with respect to the origin) exhibited by both J_{CM} and J_{A} cost surfaces.

In these ideal situations, all MSE and CM minima attain costs of zero (see Figs. 7 and 8). In addition, it can be seen that the locations of the J_{CM} and J_{A} minima coincide. (The J_{CM} minima locations can be inferred from the J_{CM} cost contours.) Fig. 9 also indicates that the curvatures of CM and MSE cost surfaces in the neighborhoods of local minima are closely related.

2) *Combined Channel-Equalizer Space:* The behavior of a gradient descent of J_{CM} is sometimes studied in the (downsampled) combined channel-equalizer space (i.e., \mathbf{h} from Section I-C). The appeal of studying J_{CM} in \mathbf{h} -space follows from the normalization and alignment of J_{CM} with the coordinate axes. These features are clear in a comparison of Figs. 9 and 10, both constructed from the same noiseless channel. Equation (4) implies that a unique reversible mapping (i.e., an isomorphism) exists between

⁶We note that in the complex-valued CM criterion, each pair of minima would be replaced by a continuum of minima spanning the full range ($0-2\pi$) of allowable system phase.

⁷Well-behaved indicates the absence of common or nearly common subchannel roots.

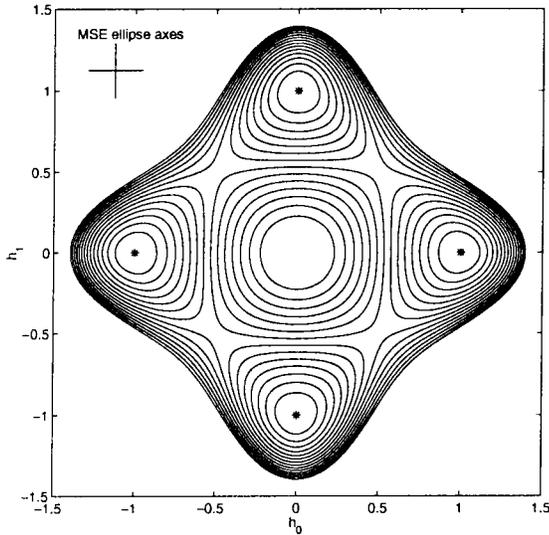


Fig. 10. J_{CM} contours for well-behaved channel c_2 and no noise in combined channel-equalizer (\mathbf{h}) space.

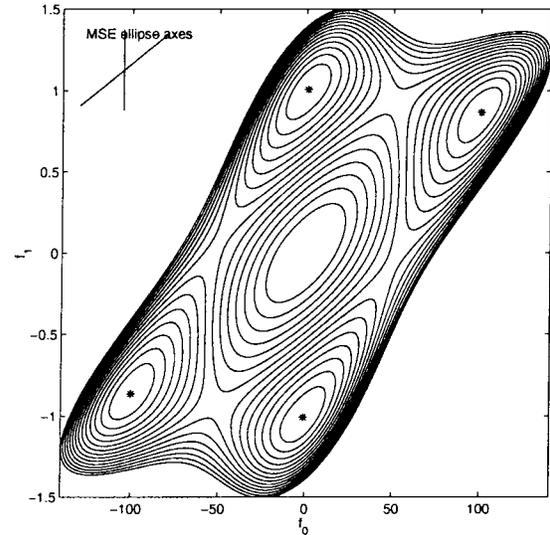


Fig. 12. J_{CM} contours for nearly common subchannel-roots channel c_3 and no noise. Note axis scaling.

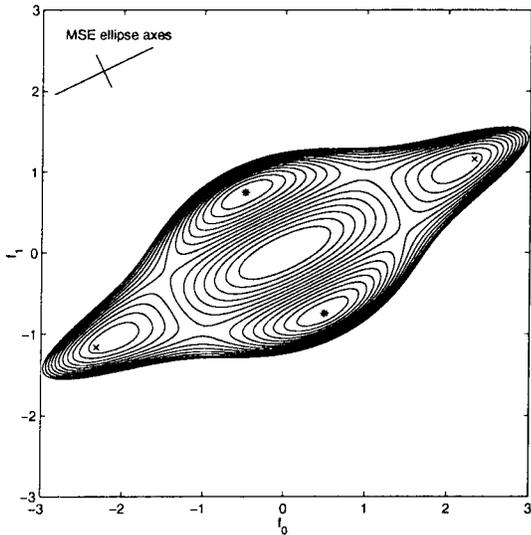


Fig. 11. J_{CM} contours for well-behaved channel c_2 and 20 dB SNR.

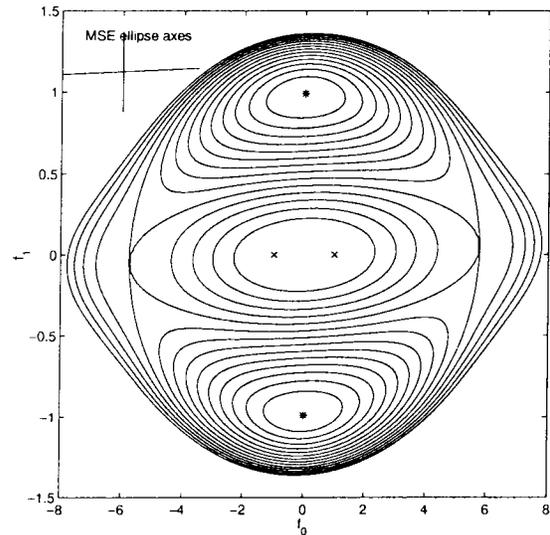


Fig. 13. J_{CM} contours for nearly common subchannel-roots channel c_3 and 20 dB SNR. Note axis scaling.

points on the J_{CM} surfaces in \mathbf{h} - and \mathbf{f} -spaces when \mathbf{C} is invertible, as it is here in our two-tap example.

3) *Additive White Channel Noise:* As channel noise is introduced, Fig. 11 indicates that the MSE and CM minima both move toward the origin in \mathbf{f} -space. The J_A and J_{CM} minima move by different amounts, though, destroying the equivalence that existed between them in the ideal case of Fig. 9. However, the relative proximity of J_A and J_{CM} minima evident in Fig. 11 still prompts consideration of J_{CM} as a close proxy for the amalgamated MSE cost J_A even when in the presence of channel noise.

4) *Common Subchannel Roots:* As evidenced by the expression we derived for MSE minima

$$\mathbf{f}^\dagger = (\mathbf{C}^H \mathbf{C} + \lambda \mathbf{I})^{-1} \mathbf{C}^H \mathbf{h}_\delta$$

when $\mathbf{C}^H \mathbf{C}$ has a large condition number, modest values of

λ can have significant consequences on \mathbf{f}^\dagger (and thus on the J_A cost surface). If the two subchannels ($c_0 + c_2 z^{-1}$ and $c_1 + c_3 z^{-1}$) have a nearly common root ($c_3/c_1 \approx c_2/c_0$) then (23) indicates that the column space of \mathbf{C} collapses; thus we expect that one eigenvalue of $\mathbf{C}^H \mathbf{C}$ will be near zero [20]. Figs. 12 and 13 use channel c_3 to demonstrate the cost surface sensitivity to noise in the presence of nearly common subchannel roots. Even under such severe surface deformation, we note that the global J_{CM} minima remain in the vicinity of global J_A minima. This further demonstrates the robustness of the relationship between J_{CM} and J_A .

5) *Channel Undermodeling:* In general, under violation of the length condition discussed in Section I-C, no equalizer settings are capable of achieving zero MSE or CM cost. This can be confirmed by extending the length of impulse response c_2 by two samples, thus forming the “undermodeled” channel c_4 . (Note that the two extra coef-

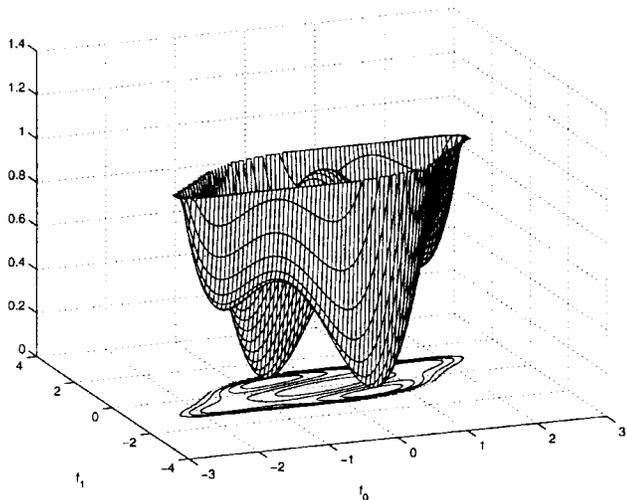


Fig. 14. J_{CM} for undermodeled channel c_4 and no noise.

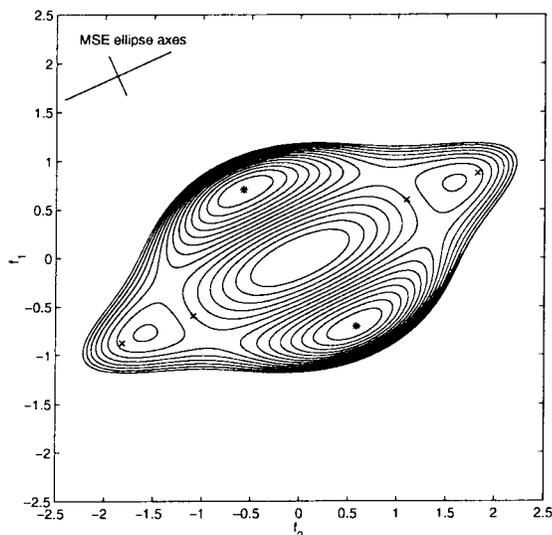


Fig. 15. J_{CM} contours for undermodeled channel c_4 and no noise.

ficients forming c_4 are no larger than any of the coefficients in c_2 .) Fig. 14 shows the CM cost surface for this undermodeled channel. Large differences in the heights of local minima demonstrate that the CM cost surface can indeed be significantly multimodal.

Elongating the channel impulse response adds another possibility for the system delay δ and thus increases the number of J_A minima (see Fig. 15). Note, however, that the number of CM minima have not changed. More importantly, note that the global CM minima remain close to their MSE counterparts under violations of the length condition.

6) *Non-CM Source:* The CM source property leading to the ideal zero-cost situation in Figs. 8–10 is violated in constructing the cost surface in Fig. 16. Here, the source is real-valued 32-PAM, which is far from CM. The non-CM property increases the source kurtosis κ_s [defined in (50)] and increases the minimum CM cost relative to that of a CM source. Notice also that the CM cost surface has become “flattened” in the parameter plane. However, as the

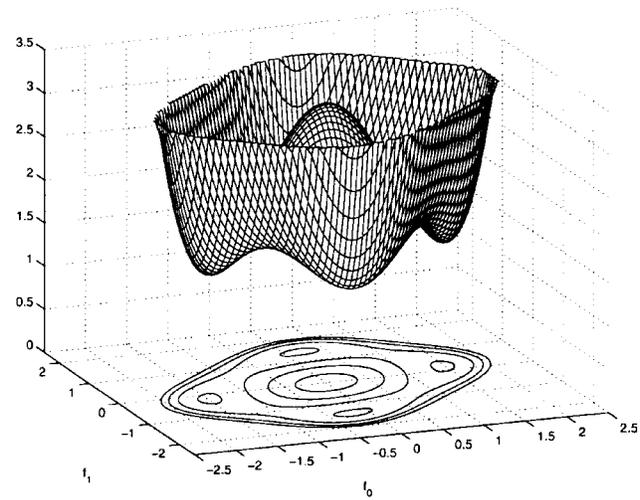


Fig. 16. Effect of source shaping ($\kappa_s = 1.8$) on J_{CM} for channel c_1 in equalizer space with no noise.

CM surface deforms due to a non-CM source, the minima locations remain unchanged.

D. Summary

Our investigations of low-dimensional examples under the following “ideal, zero-cost” conditions:

- no channel noise (i.e., $\lambda = 0$);
- no common subchannel roots (i.e., avoidance of $c_2/c_0 = c_3/c_1$);
- sufficient equalizer length (i.e., $N \geq M - 1$ for $T/2$ -spaced FSE’s);
- i.i.d., zero-mean, constant-modulus source (circularly symmetric when complex)

showed that, under such conditions, the J_A and J_{CM} minima coincide and achieve zero respective cost. Our other examples suggest that modest deviations from the ideal conditions can be tolerated in the following sense: under suitable choice of initialization, a stochastic-gradient minimization of J_{CM} will approximate the performance achieved by the same minimization of J_A . We did find, however, that the deformations caused by various violations of the ideal zero-cost conditions are different. In fact, substantial effort has been expended to characterize the performance robustness properties of the CM criterion (as descended by popular gradient descent strategies). Section III catalogs much of this effort.

The previous examples can be used to illustrate and interpret the following observations.

- *Channel noise:* CMA-based blind equalization is typically successful in common noise environments (i.e., $\sigma_s^2 > \sigma_n^2 > 0$). Under modest noise levels, relocation of global minima toward the origin is typically more severe than changes in surface curvature around such minima.
- *Undermodeling of channel length:* Given hardware constraints on equalizer length, residual ISI is unavoidable in practice. Mild contributions from un-

compensated portions of channel response typically result in mild surface deformation.

- *Nearly common subchannel roots*: These seem quite likely as channel length increases (see Fig. 18). Nearly common subchannel roots increase sensitivity to other violations from ideal conditions, but only for sub-optimal CM solutions; global CM minima still exhibit robust performance.
- *Source kurtosis*: Nonuniform (i.e., shaped) symbol distributions often lead to increased source kurtosis. As source kurtosis approaches Gaussian,⁸ the surface lifts and flattens. Lifting increases the excess error of stochastic adaptation (e.g., CMA), while flattening reduces its convergence rate. If the source exhibits a Gaussian kurtosis, the minima and saddle points vanish along a rim of the CM surface so that the gradient has solely a radial component. In this case, convergence to desirable settings is practically impossible.
- *Source correlation*: This may occur, e.g., as a result of differential encoding. Small amounts result in slight cost surface deformation. Large amounts cause major problems, such as additional local minima with terrible performance.
- *Non-CM source*: This property is unavoidable in communication systems using multilevel constellations. Though non-CM sources do not alter the minima locations, they raise and flatten the CM surface (as a consequence of increased source kurtosis—see above).
- *Initialization*: The CM surface is unavoidably multimodal. Choice of initialization affects both time-to-convergence and steady-state performance. One approach referred to in the literature suggests initializing the equalizer with a single spike⁹ time-aligned with the channel response's center of mass. In this way, crude knowledge of the channel impulse response envelope can be used to aid initialization.
- *Channel time-variation*: We proceed under the global assumption that the channel varies slowly enough in time to be tracked by the CM-minimizing gradient descent algorithm. In the vicinity of a local minimum, the tracking capabilities of any gradient descent scheme can be related to the local curvature.
- *Equalizer tap spacing*: Fractionally spaced equalizers have the ability to perfectly cancel ISI caused by a finite-length channel impulse response. In contrast, a baud-spaced equalizer requires an infinite number of taps for the same capability. Though we admit that this noiseless FIR channel model is rather academic, practical experience offers much evidence for the superiority of fractionally spaced equalization [5].
- *Transient versus steady-state performance*: Dynamic system design is often a tradeoff between transient and steady-state performance. Convergence rate is

a transient behavior descriptor; slow convergence is undesired. Excess error (due to a nonvanishing step-size and a nonzero local minimum) is a steady-state feature; abundance of excess error is undesired.

III. CM-MINIMIZING EQUALIZATION LITERATURE CATEGORIZATION

Section II presented a tutorial view of the linear equalizer design task and related the minimization of the delay-optimized and phase-indifferent mean-squared recovery error (J_A) to minimization of the CM criterion (J_{CM}). Appendix III presents a bibliography of the literature dealing with the CM criterion and its optimization via steepest gradient descent (such as with CMA). The purpose of this section is to describe our classification scheme in terms of the problem formulation and the examples of the preceding section. We also take this opportunity to cite certain papers as recommended reading on particular topics.

In addition to the birth of the CM criterion in the early 1980's, highlights in its analytical history include:

- the establishment of “perfect” conditions under which a gradient descent of the CM cost surface results in asymptotically perfect symbol recovery, i.e., “global convergence”;
- confirmation that, under slightly imperfect conditions, the CM minima remain in the vicinity of the MSE minima for various choices of delay and sign;
- recognition that, due to performance differences between CM minima under less-than-perfect conditions, initialization may be critical to acceptable transient and steady-state behavior.

The “perfect” global-convergence conditions referred to in these statements differ in detail between the baud- and fractionally spaced cases. As discussed in Section I-C, achievement of perfect source recovery devolves into exact solution of a set of simultaneous linear equations when channel noise is absent. Solution of these equations ensures that the transfer function characterizing the baud-spaced system (relating source symbols to equalized soft decisions) achieves that of a pure delay. One requirement on the existence of this perfectly equalizing solution is that the equalizer must have enough degrees of freedom. For a baud-spaced equalizer and an FIR channel, this latter requirement necessitates an equalizer with infinite impulse response (IIR) [31]. For $T/2$ -spaced FSE's, on the other hand, an equalizer response length matching (or exceeding) that of the channel proves sufficient [21]. The other requirement for the existence of a perfectly equalizing solution is that the system of equations be well posed. We mean, in an algebraic sense, that the matrix characterizing the linear system of equations must be nonsingular. For baud-spaced equalizers, this nonsingularity condition prohibits nulls in the channel frequency response (which implies, for example, that no FIR channel zeros are tolerated on the unit circle). We henceforth refer to satisfaction of this baud-spaced condition as “invertibility.” For $T/2$ -spaced FSE's, this nonsingularity translates into a

⁸Table 2 presents the values of normalized kurtosis for various sources.

⁹The single-spike initialization has its origins in baud-spaced equalization. Fractionally spaced counterparts are discussed in Section III-B3.

lack of common subchannel roots (see Appendix I-C) and is commonly referred to as “subchannel disparity.”

If conditions on the source (e.g., zero-mean, circularly symmetric, white, and sub-Gaussian) are added onto the perfect equalization requirements described in the last paragraph, a gradient descent of the CM criterion will provide asymptotically perfect source recovery from any baud- or fractionally spaced equalizer initialization. In this case the multiple CM minima all have the same depth, i.e., that of an egg carton. The distinctions in global convergence conditions between the baud- and fractionally spaced cases prompt our separation of these two cases. We note that, while analyses of CM-minimizing baud-spaced equalizers have been published since their introduction in 1980, very few analyses of CM-minimizing fractionally spaced equalizers were published before 1990.

The stringency of the global convergence requirements has prompted theoreticians to examine the impact of their violation. For example, what if the FSE length is less than the total channel response but greater than the “significant” portion of the channel response? How are prominent features of the CM cost surface (e.g., stationary point locations, regions of attraction, and heights of local minima) altered as the source is shaped or correlated and/or channel noise power increases and/or channel disparity is lost? While engineering practice desires answers about simultaneous dissatisfaction of all global convergence conditions, theoretical analysis is more likely to move forward by studying individual (or possibly pairwise) violation of these conditions. Therefore, we are encouraged to adopt a set of literature categorizations concerning studies of robustness to violations in each of the four global-convergence conditions (i.e., absence of channel noise, sufficient length, adequate disparity, and use of a zero-mean, white, circular, sub-Gaussian source process).

In Section II-C we noted that the CM and MSE error surfaces are quite similar in the vicinity of the CM local minima. This relationship implies that the local behaviors of their stochastic gradient descent minimizers (e.g., CMA and LMS, respectively) should be closely related. As a result, we are encouraged to use key behavioral descriptors associated with classical trained-LMS equalization theory as further categories for our literature classification. In particular, we borrow excess MSE (i.e., misadjustment¹⁰) and convergence rate.

While the CM and MSE criterion are comparable in a local context, their global characteristics are strikingly different. Recall the multimodality of the CM cost surface (e.g., see Figs. 8 and 14). As noted earlier, a good gradient-descent initialization may be necessary to ensure convergence to a “good” local minimum as well as to avoid temporary local capture by saddle points. In contrast, consider the trained-LMS cost surface: a unimodal elliptical hyper-paraboloid. Its unimodality obviates the need for a clever initialization strategy (assuming the training delay has been chosen). In fact, the LMS equalizer is often

¹⁰Misadjustment is defined as the ratio of excess MSE to minimum MSE.

initialized by zeroing the parameters.¹¹ If we consider delay-selection as part of the initialization of trained LMS, however, we find many similarities with the equalizer parameter initialization of CMA. Specifically, the choice of training delay bounds asymptotic LMS performance, and, in conjunction with the equalizer initialization, LMS time-to-convergence. Conversely, CMA equalizer initialization determines (asymptotic) system delay. With these thoughts in mind, we add surface topology and initialization strategy as literature categories under the heading of gradient descent behavior.

In summary, the classification scheme we adopt for our literature review uses a total of 11 labels within the three main categories discussed above.

1) Equalizer tap spacing:

- B** baud-spaced;
- F** fractionally spaced.

2) Global convergence criteria dissatisfaction:

- P** perfect; no noise, sufficient length, adequate disparity/invertibility, and zero-mean, white, circular sub-Gaussian source;
- N** noise present;
- L** equalizer length inadequate;
- D** disparity/invertibility lost or threatened;
- S** source shaped or correlated.

3) Gradient descent algorithm behavior:

- E** excess error (due to nonvanishing step-size);
- R** rate of convergence;
- T** topology of cost surface;
- I** initialization strategy.

The remainder of this section is organized by the categorization above. Each of the 11 labels is discussed using selected citations drawn from Appendix III.

Because the focus of this paper is the CM criterion in a blind linear equalizer application, we have not considered work that

- 1) principally deals with algorithm modifications (e.g., normalized, least-squares, Newton-based, block, anchored, or signed CMA) that may alter the (effective) cost function surface shape;
- 2) infers behavior principally from simulation studies with no connection made to the CM cost function;
- 3) principally addresses applications other than linear equalization (e.g., beamforming, source separation, interference cancellation, channel identification, depolarization, or decision-feedback equalization).

Though some of our citations do involve the categories above, we have chosen to include them because they contain a substantial amount of directly relevant material as well.

¹¹Initializing CMA at the origin is unwise due to the zero-valued CM-cost gradient there.

We do not provide a synopsis of each citation in the bibliography. Rather, we propose the abstracts of each paper as a source for synopses and provide a postscript bibliography that includes abstracts at http://backhoe.ee.cornell.edu/BERG/bib/CM_bib.ps.

A. Equalizer Tap Spacing

Practically speaking, the equalizer tap spacing refers to the rate at which the received signal is sampled and processed by the equalizer. In creating a discrete linear system model, the tap spacing determines the delay time of the equalizer difference equation. Using T to denote the source symbol interval, baud- or T -spaced FIR equalizers use a unit delay of T seconds in their tapped delay line. Fractionally spaced equalizers use a tap spacing less than T . The most common fractional tap spacing is $T/2$ s. In the bibliography in Section V, approximately two-thirds of the citations cover baud-spaced equalization, while the remaining one-third cover fractionally spaced equalization.

1) *Baud-Spaced Equalization*: The pioneering paper introducing the CM criterion for a complex-valued source [29] considers baud-spaced equalization only.

Conditions assuring global convergence of a baud-spaced equalizer updated via CMA are: i) no channel noise, ii) infinite impulse response equalizer, iii) no nulls in channel frequency response (i.e., no FIR channel zeros on the unit circle), and iv) a zero-mean, independent (and circularly symmetric if complex-valued) finite-alphabet source with sub-Gaussian kurtosis.

The first proof of global convergence for CMA in adapting a baud-spaced equalizer relied on a doubly infinite equalizer parameterization which allowed any combined channel-equalizer impulse response [31]. This allows convergence study in the combined channel-equalizer space, which has analytical advantages.

2) *Fractionally Spaced Equalization*: Original motivations for the use of fractional rather than baud spacing included: insensitivity to sampling phase; ability to function as a matched filter; ability to compensate for severe band-edge delay distortion; and reduced noise enhancement [5]. Fractionally spaced equalizers have nearly dominated practice since the 1980's [28]. One feature of fractionally spaced equalizers—virtually unnoticed until the 1990's—was the possibility that under ideal conditions a fractionally spaced equalizer of finite time-span could perfectly equalize an FIR channel [1]. As noted in [21], this suggests the same connection of equalizer parameters to the combined channel-equalizer parameters exploited in [31] and therefore confirms the potential for global convergence of a CM-minimizing fractionally spaced equalizer.

Conditions assuring global convergence of a $T/2$ -spaced FSE updated each baud interval via CMA are: i) no channel noise, ii) equalizer time span matching or exceeding that of the FIR channel, iii) no reflected zeros in the $T/2$ -sampled FIR channel transfer function, and iv) a zero-mean, independent (and circularly symmetric if complex-valued) finite-alphabet source with sub-Gaussian kurtosis.

These global convergence-inducing conditions do not include restriction to a constant modulus source, which was included among the ideal zero-cost conditions of Section II-D.

The first global convergence proofs for fractionally spaced CMA which do not simply rely on the extension of the baud-spaced arguments in [31] appear in [32].

B. Gradient Descent Algorithm Behavior Theory

The algorithm that performs a stochastic gradient descent of J_{CM} is often referred to as CMA

$$\mathbf{f}_{n+1} = \mathbf{f}_n + \mu \mathbf{r}_n^* y_n (\gamma - |y_n|^2). \quad (24)$$

Equation (24) is written in terms of the (fractionally sampled) regressor vector at time n

$$\mathbf{r}_n = [r_n^{\text{odd}}, \dots, r_{n-(N-1)}^{\text{odd}}, r_n^{\text{even}}, \dots, r_{n-(N-1)}^{\text{even}}]^t. \quad (25)$$

the equalizer parameter vector \mathbf{f}_n at time index n , the equalizer output y_n , a step-size μ , and the squared source-modulus γ (also referred to as the dispersion constant).

The study of dynamic systems, such as CMA, is often divided into transient and steady-state stages. Convergence rate is the dominant transient performance descriptor in classical LMS theory. MMSE and excess MSE (and their dimensionless ratio, misadjustment = EMSE/MMSE) are the dominant steady-state performance descriptors. Therefore, we consider their CM counterparts here.

Though initialization is not a major concern for the unimodal cost functions of MSE-minimizing equalizers (with preselected delay and phase), it is an unavoidable issue for CM-minimizing equalizers due to the multimodal topology of their associated cost surface. Though initialization strategies exist, none have been proven 100% successful in practice.

1) *Convergence Rate*: For trained LMS, the convergence rate (or geometric decay factor) of the sum-squared parameter error (and squared recovery error) is approximately bounded above and below by one minus twice the product of the step-size and the smallest and largest eigenvalues, respectively, of the received-signal's autocovariance matrix (i.e., $1 - 2\mu\lambda_{\min} > 1/\tau > 1 - 2\mu\lambda_{\max}$). This arises because the underlying quadratic cost function has the same Hessian, or curvature, across its entire surface. In contrast, the multimodal CM cost function has a Hessian that varies across its surface. Early convergence rate studies addressed this variation in convergence rate across the CM cost surface by focusing on convergence rate descriptors in various regions, such as far from minima and near minima [34].

Referring to Fig. 9, initialization near $[f_0, f_1] = [2.5, 0]$ will lead to a small-stepsize gradient-descent trajectory that passes through the neighborhood of a saddle point. An example displaying multiple temporary saddle-captures appears in [35]. We believe this saddle capture phenomenon to be the source of the folklore that considers CMA to be slowly converging.

A lower bound on the initialization-independent convergence rate is impossible with the multimodal CM surface due to potential of indefinite-term capture by saddle points.

In the neighborhood of a local minimum, the curvature of CMA's cost surface can be directly related to that of trained-LMS [36]. Thus, the LMS convergence rate expression can be used in a traditional manner (e.g., [23]) to provide limits on the channel tracking¹² capabilities of CMA.

2) *Excess Cost at Convergence:* In realistic situations, it is impossible to zero the update of a nonvanishing-step-size stochastic gradient descent algorithm, even at the optimum solution. With trained LMS or CMA, this undying perturbation may be a result of channel noise or residual ISI. With CMA, the nonzero update may also be the result of a non-CM source. The effect of a nonvanishing equalizer update is an asymptotic MSE level higher than that attained by the optimum fixed equalizer. This is directly related to the lifting effect that a non-CM source has on the CM cost surface, which is evident in Fig. 16.

In addition to the factors determining the excess MSE of trained LMS (i.e., stepsize, minimum achievable cost, equalizer length, and received signal power) CMA also has a term dependent on the source kurtosis.

Excess MSE of fixed (small) step-size CMA due to a non-CM source is analyzed in [37].

Figs. 8 and 16 show the effect of changing the source from constant to nonconstant modulus while simultaneously satisfying all of the global convergence conditions. Though the CM minima rise in height, they remain in the same locations in the equalizer parameter plane. As long as the source is kept sub-Gaussian, a (pure) gradient descent algorithm would still be able to asymptotically achieve perfect symbol recovery.

3) *Initialization:* As noted in the examples of Section II-C and illustrated in Figs. 11 and 15, the presence of noise or channel undermodeling causes some CM minima to achieve better performance than others.

Under violation of the conditions ensuring global convergence, choice of initialization determines asymptotic performance.

Two initialization strategies are common in the literature and in practice: spike based or matched filter. The single-spike initialization promoted in [29] for baud-spaced CMA is characterized by one nonzero equalizer tap, usually located somewhere in the central portion of the equalizer tapped-delay line. For $T/2$ -spaced CMA, a suitable extension of the single-spike idea might be a "double-spike" initialization, whereby two adjacent taps are initialized nonzero. In the frequency domain, double-spike initialization has a lowpass characteristic, a property also shared by the transmitter's pulse-shaping filter. In a mild-ISI

¹²In many practical implementations, such as those with low ambient noise levels, CMA lowers the symbol error rate to a level suitable for decision-directed LMS (DD-LMS) to take over. Due to its lower excess error, DD-LMS is preferred for tracking the slow channel variations. In low-SNR situations, however, such as those that may arise with a coded system, the tracking ability of CMA might prove important due to the potential infeasibility of DD-LMS.

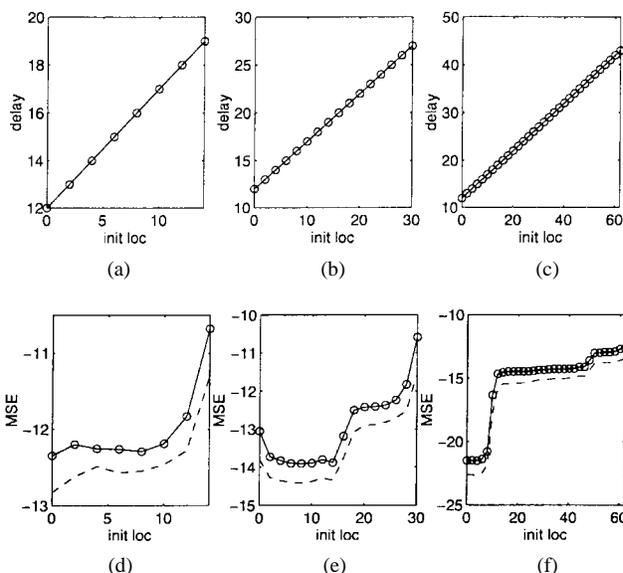


Fig. 17. (a)–(c) CMA's achieved system delay as a function of double-spike location for an equalizer with length 16, 32, 64, respectively. (d)–(f) CMA's asymptotic MSE performance (solid) compared to same-delay MMSE performance (dashed) for an equalizer with length 16, 32, 64, respectively.

environment, one might even consider initializing the FSE with an impulse response matching the pulse-shaping filter itself, as (in this mild case) this response is close to the expected steady-state equalizer solution (assuming that the FSE is used to accomplish matched filtering at the receiver).

All of the initialization techniques above still require a selection of delay, i.e., spike positioning within the equalizer time span. This delay choice is intimately connected to the delay-choice in trained-LMS equalization in the following way: CMA tends to converge to minima with the same group delay as its initialization. Fig. 17 provides evidence for this claim using double-spike initializations of $T/2$ -spaced CMA on the SPIB microwave channel shown in Fig. 1 under 50 dB SNR and a QPSK source. Note the (affine) linear correspondence between double-spike position and asymptotically achieved system delay. Another interesting characteristic of Fig. 17, seen after comparing subplots Fig. 17(d)–(f) to Fig. 6, is its suggestion that *the set of system delays reachable by CMA are best in an MMSE sense*. We offer these last two statements as educated conjectures, as no theoretical proofs yet exist to verify them.

The aforementioned relationship between initialization and channel group delays suggests that prior information about the channel may aid in selection of initialization delay choice. Appendix III notes the existence of other more complicated off-line initialization schemes that leverage such notions.

4) *Surface Topology:* In Figs. 8 and 14, the "molar" shape of the CM cost surface in two-tap real-valued equalizer space is the same one used in Section II-C to aid in an understanding of CMA's transient and asymptotic performance, as well as to motivate the importance of

initialization. Section II-C also described how deformation of this molar shape occurs with violation of the various ideal zero-cost conditions, and it used this surface-centric view to predict the pertinent effects of these violations.

The three-dimensional “molar” shape typical of the real-valued two-tap equalizer CM cost surface offers a compact visualization of virtually all of the major features of CMA behavior theory, applicable even to longer equalizers.

Surface characterization via gradient and Hessian formulas is provided in [38] for baud-spaced equalizers. Reference [39] offers a more developed topological study of the fractionally spaced CM criterion.

C. Violation of Conditions Ensuring Global Convergence

1) *Perfect—All Conditions Satisfied:* While Sections III-A1 and III-A2 listed conditions ensuring the global convergence of CMA, their violation is unavoidable in practice.

There exists a set of conditions under which an arbitrarily initialized gradient-descent minimization of the CM criterion results in perfect symbol recovery. These “global convergence” inducing conditions, however, are unconditionally violated—if only modestly—in practice.

2) *Channel Noise Present:* CM-based blind equalization typically remains successful in common noise environments (i.e., $\sigma_s^2 > \sigma_n^2 > 0$). To recall the cost surface deformations due to noise, compare Figs. 9–11.

When the presence of (modest) channel noise is the only violation of the global convergence conditions, the locations of global CM minima shift toward the origin in equalizer parameter space and the minimum achievable CM cost is increased.

This behavior is strikingly similar to the behavior of the MSE criterion in the presence of channel noise. In fact, under modest amounts of noise, the CM minima remain near the MSE minima [33], [40].

At extremely high noise levels (i.e., $\sigma_n^2 > \sigma_s^2$), the two criteria differ in the following manner: the MSE minima continue to move toward the origin, while the CM minima remain within an annulus outside the origin. This behavior is attributed to the so-called “CMA power constraint” [40].

We have also observed the disappearance of local minima under modest-to-high noise levels [41], especially for channels without much disparity (see Fig. 13).

3) *Insufficient Equalizer Length:* In order to completely cancel the ISI induced by an arbitrary FIR channel, one requires an IIR baud-spaced equalizer or a sufficiently long FIR fractionally spaced equalizer. In the presence of channel noise, the MSE-optimal equalizer makes a compromise between ISI cancellation and noise gain, and the resulting equalizer impulse response is no longer finite-length, even for fractionally spaced equalizers [5].

In the presence of noise, the (baud- and fractionally spaced) MMSE equalizers have an infinite impulse response, implying that the length of an FIR equalizer should be chosen to capture “enough” of the desired response.

Studies on the effect of violations in the equalizer length condition include [42] in a baud-spaced context, and [43], [44] in a fractionally spaced context. The latter provide

evidence of CMA robustness to modest channel undermodeling and include approximate bounds on performance.

As hardware advances permit-increased baud-rate, physical channel delay-spreads remain unchanged, and the relative length of the channel impulse response grows proportionally. To combat ISI, there is a corresponding need to increase equalizer length. Therefore, the desire for higher communication rates will always stress the equalization task. This is a primary justification for the continued development and study of truly simple adaptive equalization algorithms like LMS and CMA.

4) *Disparity/Invertibility Lost:* As discussed earlier, the set of zero-forcing equalizer design equations becomes poorly conditioned in the presence of deep spectral nulls for baud-spaced equalizers, or in the presence of nearly common subchannel roots for fractionally spaced equalizers. Poor conditioning implies an increased parameter sensitivity to noise and other violations of the global convergence conditions. Fortunately, this parameter sensitivity does not imply a performance sensitivity. In other words, global CMA minima remain robust under a loss of disparity. We note that the same is true for the delay-optimal MMSE solutions.

A near-loss of disparity (for FSE’s) or invertibility (for BSE’s) dramatically increases the sensitivity of suboptimal CM (and MSE) minima to other violations in the global convergence conditions. However, global CM (and MSE) minima remain robust under these conditions.

The behavior of fractionally spaced CM (and MSE) minima under loss of disparity is explained through the following design procedure. For simplicity, let us assume the absence of noise. 1) Factor the common root(s) out of the subchannels in Fig. 5 and form a new system composed of the common root(s) component and what remains of the multichannel component, connected in series. 2) Design the subequalizers so that the remaining multichannel component approximates the inverse of the common root(s) component. At this point, the cascaded system should approximate a pure delay. This procedure closely describes the construction of the MMSE or CM-optimal equalizers under a loss of disparity [33]. We describe this idea more formally in Appendix I-C.

There are a number of reasons that we expect the presence of nearly common subchannel roots, e.g., nearly reflected¹³ $T/2$ -spaced roots, in realistic situations. Looking at Fig. 18, which portrays the roots of the length 300 $T/2$ -sampled SPIB channel whose impulse response appears in Fig. 1 and whose response we consider to be “typical,” one notices the apparent plethora of nearly reflected roots. Similarly, one might realize that a long FIR approximation to a pole¹⁴ in the physical channel would also generate nearly reflected roots. These reasons suggest the likelihood

¹³Common subchannel roots have been shown to be identical to $T/2$ -spaced channel roots reflected across the origin [25].

¹⁴A degree- N polynomial forming a close approximation to a single pole can be constructed using N roots on a ring in the complex plane with a radius equal to the pole magnitude. The roots are spaced at $N + 1$ equal intervals on the ring with the exception that there exists no root at the location of the approximated pole.

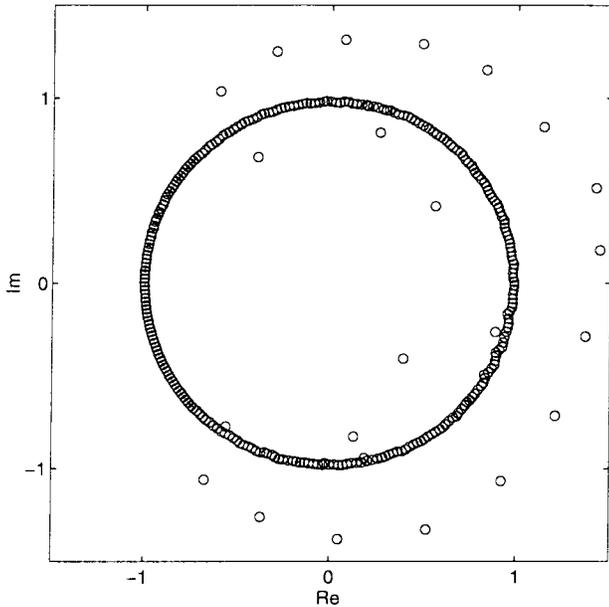


Fig. 18. Roots of $T/2$ -sampled SPIB terrestrial microwave channel #3.

of nearly common subchannel roots in realistic situations. See [2] for further discussion on the existence of reflected roots in physical systems and its negative implications on second-order-statistics based blind equalization.

5) *Shaped or Correlated Source*: Source shaping, encouraged by a potential increase in coding gain (e.g., see [3]), has the effect of making the source symbol distribution more Gaussian. As far as our problem is concerned, it has the practical effect of raising the kurtosis. Increases in source kurtosis, as long as they remain sub-Gaussian, do not affect the locations of CM local minima. However, they are known to flatten the CM cost surface in all but the radial direction, making CMA's convergence to the minima slower (and in the limiting Gaussian case, impossible). In addition, increases in source kurtosis have been shown to raise the CM surface (see Fig. 16), thus increasing the excess asymptotic error levels achieved by nonvanishing-step-size stochastic gradient algorithms.

Recall that non-CM sources also have kurtoses greater than one. To put source shaping in perspective, Table 2 presents the kurtosis of popular source alphabets along with the limiting Gaussian values. Note that a shaped source has the potential for raising the kurtosis far past that of a dense (uniform) constellation like 1024-QAM.

For shaped sources with near-Gaussian kurtoses, the CM cost surface is raised and flattened, therefore unsuited to stochastic gradient descent.

Source correlation results from the use of certain types of coding (e.g., differential encoding) or under particular operational circumstances [45], [46]. Moderate amounts of source correlation may shift the locations of local minima. Large amounts of correlation may cause additional (false) minima to appear in the CM cost surface. Recall that any amount of source correlation violates the CM global convergence requirements. The most thorough studies on

Table 2 Normalized Kurtoses for Various Source Distributions

real-valued alphabet	kurtosis	complex-valued alphabet	kurtosis
uniform BPSK	1	uniform M -PSK	1
uniform 4-PAM	1.64	uniform 16-QAM	1.32
uniform 8-PAM	1.762	uniform 64-QAM	1.381
uniform 16-PAM	1.791	uniform 256-QAM	1.395
uniform 32-PAM	1.798	uniform 1024-QAM	1.399
Gaussian	3	Gaussian	2

Table 3 Annotations Used in the Bibliography and Their Interpretations

Symbol	Meaning	Comments
P	Perfect Equalization	Addresses the case where any pure delay is achievable; global convergence
N	Noise	Addresses the effects of noise
L	Length	Addresses the effects of equalizer length
D	Disparity	Addresses the effects of channel disparity
S	Source	Addresses the effects of sources that are shaped, non-constant modulus or correlated
F	FSE	Fractionally-spaced equalization context
B	BSE	Baud-spaced equalization context
I	Initialization	Discusses initialization procedures for adaptive implementations
E	Excess error	Discusses sources of excess error in adaptive implementations due to nonvanishing stepsize
R	Convergence Rate	Discusses convergence rate of adaptive implementations
T	Surface Topology	Studies topology of error surface

the effects of shaped and/or correlated sources appear in [39] and [46].

As a final note, we point out that the global convergence conditions for complex-valued implementations of the CM criterion specify a circularly symmetric source, i.e., $E\{s_n^2\} = 0$. Studies have shown that violations of this requirement (e.g., from the use of a real-valued source with a complex-valued channel and/or equalizer) can result in the appearance of undesired CM minima [47].

D. Additional Information

These descriptions of the literature categorization have prepared the reader to utilize the annotated bibliography in Appendix III, which provides an in-depth look into the CM literature. Each entry in the bibliography is annotated with boldface letters that indicate the classification of its content (see Table 3). A postscript file containing the abstracts of papers in this list is provided available from: http://backhoe.ee.cornell.edu/BERG/bib/CM_bib.ps.

We also recommend The BERGULATOR, a public-domain MATLAB-5 based software environment which allows for experimentation with the CM criterion and various implementations of CMA. It can be used, for example, to generate all contour plots in this paper. The BERGULATOR was written by P. Schniter of Cornell University's Blind Equalization Research Group (CU-BERG) and is available from the following web site: <http://backhoe.ee.cornell.edu/BERG/>.

We denote the combined LTI channel and pulse-shape impulse response by $c(t)$ and the baseband additive channel noise process by $w(t)$. The continuous-time baseband representation of the waveform seen by the receiver can then be described by

$$r(t) = \sum_{n=-\infty}^{\infty} s_n c(t - nT - t_0) + w(t) \quad (26)$$

for symbol sequence $\{s_n\}$, baud interval T , and arbitrary time delay t_0 . Sampling¹⁵ the received signal every $T/2$ s at the receiver, we denote the sampled received sequence by

$$r\left(k \frac{T}{2}\right) = \sum_{n=-\infty}^{\infty} s_n c\left(k \frac{T}{2} - nT - t_0\right) + w\left(k \frac{T}{2}\right). \quad (27)$$

The output x_k of a length $2N$ FIR FSE with tap spacing of $T/2$ can be written as a $T/2$ -rate convolution with the sampled received sequence

$$x_k = \sum_{i=0}^{2N-1} f_i r\left((k-i) \frac{T}{2}\right). \quad (28)$$

The choice of an even number of equalizer taps is chosen for notational simplicity. Now suppose that only the “odd” fractionally spaced equalizer output samples are retained in a decimation by two (i.e., $k = 2n + 1$ for $n = 0, 1, 2, \dots$). The decimated equalized output sequence y_n^{odd} then becomes

$$y_n^{\text{odd}} = x_{2n+1} \quad (29)$$

$$= \sum_{i=0}^{2N-1} f_i r\left(nT - i \frac{T}{2} + \frac{T}{2}\right) \quad (30)$$

$$= \sum_{i=0}^{N-1} \left(f_{2i} r\left((n-i)T + \frac{T}{2}\right) + f_{2i+1} r((n-i)T) \right). \quad (31)$$

Note that a similar procedure can be carried out for even-indexed output sampling (i.e., $k = 2n$ and $y_n^{\text{even}} = x_{2n}$). An illustration of the setup described above appears in Fig. 3.

A. Multichannel Model

From (31) we observe that the decimated output y_n^{odd} can be considered the sum of two baud-spaced convolutions

$$y_n^{\text{odd}} = \sum_{i=0}^{N-1} (f_i^{\text{even}} r_{n-i}^{\text{odd}} + f_i^{\text{odd}} r_{n-i}^{\text{even}}) \quad (32)$$

where

$$f_n^{\text{even}} = f_{2n} \quad f_n^{\text{odd}} = f_{2n+1} \quad r_n^{\text{even}} = r(nT)$$

¹⁵The noise and channel are considered band-limited assuming antialias filtering is done prior to $T/2$ -spaced sampling at the receiver.

and

$$r_n^{\text{odd}} = r\left(nT + \frac{T}{2}\right). \quad (33)$$

We refer to r_n^{even} and r_n^{odd} as the “even” and “odd” received sequences and to f_n^{even} and f_n^{odd} as the “even” and “odd” subequalizers, respectively.

Defining the even and odd baud-rate channel response samples

$$c_n^{\text{even}} = c(nT - t_0) \quad \text{and} \quad c_n^{\text{odd}} = c\left(nT + \frac{T}{2} - t_0\right) \quad (34)$$

and channel noise samples $w_n^{\text{even}} = w(2n(T/2))$ and $w_n^{\text{odd}} = w((2n+1)(T/2))$ (for nonnegative integers n), we can confirm that they are related to the received subsequences in a straightforward manner

$$r_n^{\text{even}} = \sum_l s_l c_{n-l}^{\text{even}} + w_n^{\text{even}} \quad (35)$$

$$r_n^{\text{odd}} = \sum_l s_l c_{n-l}^{\text{odd}} + w_n^{\text{odd}}. \quad (36)$$

These expressions allow us to rewrite the decimated equalizer output in terms of the baud-spaced symbol sequence.

It is important to note that the arbitrary delay t_0 has been incorporated into our definitions of the channel response samples. This implies that the “even” and “odd” subchannel classifications are merely notational and have no real physical significance. Furthermore, the inclusion of arbitrary delay implies that our convention of retaining the odd-indexed (as opposed to the even-indexed) decimated equalizer output samples also lacks practical significance. In this spirit, we drop the “odd” notation on y_n^{odd} and simply refer to the baud-spaced system output samples as y_n . Here we are seeing evidence for the inherent baud-synchronization capabilities of an FSE (not characteristic of BSE’s).

Substituting the received subsequence expressions (35) and (36) into (32)

$$\begin{aligned} y_n &= \sum_{i=0}^{N-1} f_i^{\text{even}} \left(\sum_l s_l c_{n-i-l}^{\text{odd}} + w_{n-i}^{\text{odd}} \right) \\ &+ \sum_{i=0}^{N-1} f_i^{\text{odd}} \left(\sum_l s_l c_{n-i-l}^{\text{even}} + w_{n-i}^{\text{even}} \right) \quad (37) \\ &= s_n \star (f_n^{\text{even}} \star c_n^{\text{odd}} + f_n^{\text{odd}} \star c_n^{\text{even}}) \\ &+ f_n^{\text{even}} \star w_n^{\text{odd}} + f_n^{\text{odd}} \star w_n^{\text{even}} \quad (38) \end{aligned}$$

where the “ \star ” indicates convolution. The relationships between the source, noise, subequalizers, and subchannels described above appears in the multichannel model of Fig. 5.

Consider for a moment the noiseless case. The impulse response h_n from transmitted source to baud-spaced equalizer output follows immediately from consideration of s_n as the Kronecker delta sequence δ_n . Thus, we conclude that

$$h_n = f_n^{\text{even}} \star c_n^{\text{odd}} + f_n^{\text{odd}} \star c_n^{\text{even}}. \quad (39)$$

This impulse response leads directly to a transfer function $H(z^{-1})$ with unit delay (z^{-1}) of duration T

$$H(z^{-1}) = F_{\text{even}}(z^{-1})C_{\text{odd}}(z^{-1}) + F_{\text{odd}}(z^{-1})C_{\text{even}}(z^{-1}). \quad (40)$$

Note that the perfect zero-forcing system $H(z) = z^{-\delta}$ (with nonnegative integer delay δ), leads to the Bezout relationship [8]

$$z^{-\delta} = F_{\text{even}}(z^{-1})C_{\text{odd}}(z^{-1}) + F_{\text{odd}}(z^{-1})C_{\text{even}}(z^{-1}). \quad (41)$$

B. Multirate Model

To show that the multirate model of Fig. 4 also originates from the fractionally spaced communication system of Fig. 3, we show that the fractionally spaced equalizer output $\{x_k\}$ in (28) can be written in terms of a zero-filled version of the source sequence $\{a_k\}$

$$a_k = \begin{cases} s_{k/2}, & \text{for } k \text{ even} \\ 0, & \text{for } k \text{ odd} \end{cases} \quad (42)$$

as depicted in Fig. 4. Rewriting (27) as

$$r\left(k\frac{T}{2}\right) = \sum_{l=-\infty}^{\infty} a_l c\left((k-l)\frac{T}{2} - t_0\right) + w_k \quad (43)$$

we see upon its substitution into (28) that

$$x_k = \sum_{i=0}^{2N-1} f_i \left(\sum_l a_l c\left((k-i-l)\frac{T}{2} - t_0\right) + w_{k-i} \right) \quad (44)$$

$$= \sum_{i=0}^{2N-1} f_i \left(\sum_l a_l c_{k-i-l} + w_{k-i} \right) \quad (45)$$

$$= f_k \star (a_k \star c_k + w_k) \quad (46)$$

where the fractionally spaced channel response samples c_k are defined such that $c_k = c(k(T/2) - t_0)$.

At this point we can observe that, in the noiseless case, the fractionally spaced system impulse response h_k^{FS} becomes

$$h_k^{\text{FS}} = f_k \star c_k. \quad (47)$$

Note from (39) that only half of the terms in the fractionally spaced impulse response (47) are directly relevant to the system output since the fractionally spaced output $\{x_k\}$ is later decimated by two.

C. The Subchannel Disparity Condition

The Bezout equation (41) leads directly to the perfect equalization requirement concerning subchannel roots. Specifically, for the existence of a (finite-length) zero-forcing equalizer, the subchannel polynomials $C_{\text{even}}(z^{-1})$ and $C_{\text{odd}}(z^{-1})$ must not share a common root.

The existence of perfectly equalizing sub-equalizer polynomials $F_{\text{even}}(z^{-1})$ and $F_{\text{odd}}(z^{-1})$ implies that (41) can

be satisfied. For example, if the subchannels share one root, a common polynomial $G(z^{-1}) = g_0 + g_1 z^{-1}$ can be factored out of both $C_{\text{even}}(z^{-1})$ and $C_{\text{odd}}(z^{-1})$, leaving $\bar{C}_{\text{even}}(z^{-1})$ and $\bar{C}_{\text{odd}}(z^{-1})$, respectively. The perfect equalization relationship would then become

$$z^{-\delta} = G(z^{-1}) (F_{\text{even}}(z^{-1})\bar{C}_{\text{odd}}(z^{-1}) + F_{\text{odd}}(z^{-1})\bar{C}_{\text{even}}(z^{-1})) \quad (48)$$

but this is contradicted by the fact that there is no finite-length polynomial $F_{\text{even}}(z^{-1})\bar{C}_{\text{odd}}(z^{-1}) + F_{\text{odd}}(z^{-1})\bar{C}_{\text{even}}(z^{-1})$ that when multiplied by $G(z^{-1})$ results in the delay operator $z^{-\delta}$.

However, $F_{\text{even}}(z^{-1})$ and $F_{\text{odd}}(z^{-1})$ can be chosen so that (48) is approximated, in which case the following relationship is satisfied:

$$z^{-\delta} G^{-1}(z^{-1}) \approx F_{\text{even}}(z^{-1})\bar{C}_{\text{odd}}(z^{-1}) + F_{\text{odd}}(z^{-1})\bar{C}_{\text{even}}(z^{-1}). \quad (49)$$

In other words, the FSE combines with the noncommon-root component of the channel to approximate the (IIR) inverse of the (T -spaced) common root component.

D. On The Independence of Fractionally Sampled Channel Noise

A typical assumption on the (baseband equivalent) channel noise $w(t)$ is that it is well modeled by a zero-mean, circularly symmetric Gaussian process [10]. In many situations $w(t)$ is also assumed to have a flat wideband power spectrum. Does this imply that the fractionally sampled noise process $\{w_k\}$ will also be white? Under these conditions, $\{w_k\}$ will only be white when the anti-alias filters prior to $T/2$ -spaced sampling satisfy a rate $2/T$ Nyquist criterion. In practice, this criterion is satisfied by anti-alias filters that are power-symmetric about the frequency $1/T$ Hz. If, for example, the filtering prior to equalization is matched to the pulse shape of the transmitted signal, then fractionally sampled $\{w_k\}$ will not be white.

APPENDIX II

THE CM COST FUNCTION

Below we provide the general formulation of the CM cost function for a complex i.i.d. zero-mean source and complex baseband channel in additive white zero-mean noise. We will assume that each member of the symbol alphabet is equiprobable in the source sequence. Furthermore, we also assume that the receiver sampling clock is frequency synchronous (a fixed time offset is allowed) with the source symbol clock. In practice, this is a reasonable assumption since the symbol clock can often be extracted by computing the square magnitude of the received signal (commonly known as envelope detection). Given these assumptions, we follow the general formulation of the CM cost function with expressions for the specific cases of PAM, PSK, and QAM input signals.

In addition to the previously introduced notation we will use the following definitions:

$$\kappa_s = \frac{E\{|s_n|^4\}}{\sigma_s^4}, \text{ the normalized kurtosis of } \{s_n\} \quad (50)$$

$$\gamma = \frac{E\{|s_n|^4\}}{\sigma_s^2}, \text{ the dispersion constant of } \{s_n\} \quad (51)$$

$$\|\mathbf{h}\|_2^2 = \sum_{n=0}^{P-1} |h_n|^2, \text{ the squared } \ell_2\text{-norm of } \mathbf{h}. \quad (52)$$

Note that $\gamma = \sigma_s^2 \kappa_s$, where $\sigma_s^2 = E\{|s_n|^2\}$. Following the presentation of the FS system model in Section I-C and Appendix I, we can redefine the equalizer output using (4) and (13). This results in

$$y_n = \mathbf{h}^t \mathbf{s}(n) + \mathbf{f}^t \mathbf{w}(n). \quad (53)$$

The CM cost function is

$$\begin{aligned} J_{\text{CM}} &= E\left\{(\gamma - |y_n|^2)^2\right\} \\ &= E\{|y_n|^4\} - 2\gamma E\{|y_n|^2\} + \gamma^2 \\ &= E\{|y_n|^4\} - 2\sigma_s^2 \kappa_s E\{|y_n|^2\} + \sigma_s^4 \kappa_s^2. \end{aligned} \quad (54)$$

In order to analyze J_{CM} , we will first expand $|y_n|^2$, using (53). For convenience we will temporarily let $A_n = \mathbf{h}^t \mathbf{s}(n)$ and $B_n = \mathbf{f}^t \mathbf{w}(n)$, where $y_n = A_n + B_n$. Using the assumptions of mutually independent zero-mean noise and source sequences, we note that A_n and B_n are also independent and zero-mean, i.e.,

$$\begin{aligned} E\{A_n\} &= \mathbf{h}^t E\{\mathbf{s}(n)\} = 0 \\ E\{B_n\} &= \mathbf{f}^t E\{\mathbf{w}(n)\} = 0 \end{aligned}$$

and

$$E\{A_n B_n\} = E\{A_n\} E\{B_n\}.$$

With these assumptions, we arrive at

$$\begin{aligned} E\{|y_n|^2\} &= E\{|A_n|^2\} + E\{A_n\} E\{B_n^*\} \\ &\quad + E\{A_n^*\} E\{B_n\} + E\{|B_n|^2\} \\ &= E\{|A_n|^2\} + E\{|B_n|^2\}. \end{aligned} \quad (56)$$

Expanding $|A_n|^2$ and $|B_n|^2$, we have that

$$E\{|y_n|^2\} = \sigma_s^2 \|\mathbf{h}\|_2^2 + \sigma_w^2 \|\mathbf{f}\|_2^2. \quad (57)$$

The same approach can be used to examine $E\{|y_n|^4\}$, which leads to the following equation:

$$\begin{aligned} E\{|y_n|^4\} &= E\{|A_n|^4\} + E\{A_n^2\} E\{(B_n^*)^2\} \\ &\quad + 4E\{|A_n|^2\} E\{|B_n|^2\} \\ &\quad + E\{B_n^2\} E\{(A_n^*)^2\} + E\{|B_n|^4\}. \end{aligned} \quad (58)$$

$$\quad (59)$$

Due to space limitations, we omit the details of the derivation of $E\{|y_n|^4\}$ but mention the following properties used in the derivation.

- The second-order terms are relatively easy to compute; they involve summations of source (and noise) terms of the form $E\{s_{n-i} s_{n-l}\}$, $E\{s_{n-i} s_{n-l}^*\}$, or $E\{s_{n-i}^* s_{n-l}^*\}$.

- The fourth-order terms are more difficult to compute, but each of the source (and noise) terms are of the form $E\{s_{n-i} s_{n-l}^* s_{n-m} s_{n-j}^*\}$.

Any of the expectations not involving an even power (two or four) will vanish because the source and noise are both zero-mean and white. After a considerable amount of algebra we arrive at the following expression for $E\{|y_n|^4\}$. Noting that $E\{s_n^2\}$ and $E\{w_n^2\}$ are independent of n , we will denote expectations of this form by $E\{s^2\}$ and $E\{w^2\}$, respectively

$$\begin{aligned} E\{|y_n|^4\} &= \kappa_s \sigma_s^4 \sum_{i=0}^{P-1} |h_i|^4 + 2\sigma_s^4 \sum_{i=0}^{P-1} \sum_{m=0, m \neq i}^{P-1} |h_i|^2 \\ &\quad \cdot |h_m|^2 + |E\{s^2\}|^2 \sum_{i=0}^{P-1} \sum_{j=0, j \neq i}^{P-1} h_i^2 (h_j^*)^2 \\ &\quad + \kappa_w \sigma_w^4 \sum_{i=0}^{2N-1} |f_i|^4 + 2\sigma_w^4 \sum_{i=0}^{2N-1} \sum_{m=0, m \neq i}^{2N-1} \\ &\quad \cdot |f_i|^2 |f_m|^2 + |E\{w^2\}|^2 \sum_{i=0}^{2N-1} \sum_{j=0, j \neq i}^{2N-1} f_i^2 (f_j^*)^2 \\ &\quad + \left(E\{s^2\} \sum_{i=0}^{P-1} h_i^2 \right) \left(E\{w^2\} \sum_{i=0}^{2N-1} f_i^2 \right)^* \\ &\quad + 4\sigma_s^2 \sigma_w^2 \|\mathbf{h}\|_2^2 \|\mathbf{f}\|_2^2 + \left(E\{s^2\} \sum_{i=0}^{P-1} h_i^2 \right)^* \\ &\quad \cdot \left(E\{w^2\} \sum_{i=0}^{2N-1} f_i^2 \right). \end{aligned} \quad (60)$$

We define the noise kurtosis κ_w analogous to the source kurtosis κ_s in (50). Substituting (57) and (60) into (55) we have the final expansion of the cost function.

$$\begin{aligned} J_{\text{CM}} &= \kappa_s \sigma_s^4 \sum_{i=0}^{P-1} |h_i|^4 + 2\sigma_s^4 \sum_{i=0}^{P-1} \sum_{m=0, m \neq i}^{P-1} \\ &\quad \cdot |h_i|^2 |h_m|^2 + |E\{s^2\}|^2 \sum_{i=0}^{P-1} \sum_{j=0, j \neq i}^{P-1} h_i^2 (h_j^*)^2 \\ &\quad + \kappa_w \sigma_w^4 \sum_{i=0}^{2N-1} |f_i|^4 + 2\sigma_w^4 \sum_{i=0}^{2N-1} \sum_{m=0, m \neq i}^{2N-1} \\ &\quad \cdot |f_i|^2 |f_m|^2 + |E\{w^2\}|^2 \sum_{i=0}^{2N-1} \sum_{j=0, j \neq i}^{2N-1} f_i^2 (f_j^*)^2 \\ &\quad + \left(E\{s^2\} \sum_{i=0}^{P-1} h_i^2 \right) \left(E\{w^2\} \sum_{i=0}^{2N-1} f_i^2 \right)^* \\ &\quad + 4\sigma_s^2 \sigma_w^2 \|\mathbf{h}\|_2^2 \|\mathbf{f}\|_2^2 + \left(E\{s^2\} \sum_{i=0}^{P-1} h_i^2 \right)^* \\ &\quad \cdot \left(E\{w^2\} \sum_{i=0}^{2N-1} f_i^2 \right) - 2\sigma_s^2 \kappa_s \\ &\quad \cdot (\sigma_s^2 \|\mathbf{h}\|_2^2 + \sigma_w^2 \|\mathbf{f}\|_2^2) + \sigma_s^4 \kappa_s^2. \end{aligned} \quad (61)$$

We will now consider how various restrictions on the source and noise simplify this equation.

A. PAM Source, Real-Valued Channel

For PAM the source symbols s_n are real valued so that $E\{|s_n|^2\} = E\{s_n^2\} = \sigma_s^2$. Furthermore, if w_n , \mathbf{f} , and \mathbf{h} are real valued, we have $E\{|w_n|^2\} = E\{w_n^2\} = \sigma_w^2$, $f_i^2 = (f_i^*)^2 = |f_i|^2$, and $h_i^2 = (h_i^*)^2 = |h_i|^2$. Thus, we have that, for a real-valued source and real-valued channel, (61) reduces to

$$\begin{aligned} J_{\text{CM}}|_{\text{PAM}} &= \kappa_s \sigma_s^4 \sum_{i=0}^{P-1} h_i^4 + 3\sigma_s^4 \sum_{i=0}^{P-1} \sum_{m=0, m \neq i}^{P-1} h_i^2 h_m^2 \\ &+ \kappa_w \sigma_w^4 \sum_{i=0}^{2N-1} f_i^4 + 3\sigma_w^4 \sum_{i=0}^{2N-1} \sum_{m=0, m \neq i}^{2N-1} \\ &\cdot f_i^2 f_m^2 + \sigma_s^2 \sigma_w^2 \|\mathbf{h}\|_2^2 \|\mathbf{f}\|_2^2 + 4\sigma_s^2 \sigma_w^2 \|\mathbf{h}\|_2^2 \|\mathbf{f}\|_2^2 \\ &+ \sigma_s^2 \sigma_w^2 \|\mathbf{h}\|_2^2 \|\mathbf{f}\|_2^2 \\ &- 2\sigma_s^2 \kappa_s (\sigma_s^2 \|\mathbf{h}\|_2^2 + \sigma_w^2 \|\mathbf{f}\|_2^2) + \sigma_s^4 \kappa_s^2. \end{aligned}$$

Noting that

$$\sum_{i=0}^{P-1} \sum_{m=0, m \neq i}^{P-1} h_i^2 h_m^2 = \|\mathbf{h}\|_2^4 - \sum_{i=0}^{P-1} h_i^4$$

and summing like terms we arrive at

$$\begin{aligned} J_{\text{CM}}|_{\text{PAM}} &= \sigma_s^4 (\kappa_s - 3) \sum_{i=0}^{P-1} h_i^4 + 3\sigma_s^4 \|\mathbf{h}\|_2^4 \\ &+ \sigma_w^4 (\kappa_w - 3) \sum_{i=0}^{2N-1} f_i^4 + 3\sigma_w^4 \|\mathbf{f}\|_2^4 \\ &+ 6\sigma_s^2 \sigma_w^2 \|\mathbf{h}\|_2^2 \|\mathbf{f}\|_2^2 \\ &- 2\sigma_s^2 \kappa_s (\sigma_s^2 \|\mathbf{h}\|_2^2 + \sigma_w^2 \|\mathbf{f}\|_2^2) + \sigma_s^4 \kappa_s^2. \end{aligned} \quad (62)$$

Note that if the noise is Gaussian, $\kappa_w = 3$ and the third term in (62) is zero.

1) *BPSK Source, Real-Valued Channel*: Here we consider the subcase of a BPSK source in a real-valued channel results in further simplifications. For BPSK, $\kappa_s = \sigma_s^4 = 1$, which implies that (62) reduces to

$$\begin{aligned} J_{\text{CM}}|_{\text{BPSK}} &= -2 \sum_{i=0}^{P-1} h_i^4 + 3\|\mathbf{h}\|_2^4 + \sigma_w^4 (\kappa_w - 3) \\ &\cdot \sum_{i=0}^{2N-1} f_i^4 + 3\sigma_w^4 \|\mathbf{f}\|_2^4 + 6\sigma_w^2 \|\mathbf{h}\|_2^2 \|\mathbf{f}\|_2^2 \\ &- 2(\|\mathbf{h}\|_2^2 + \sigma_w^2 \|\mathbf{f}\|_2^2) + 1. \end{aligned} \quad (63)$$

In the absence of noise, (63) is the equation given in [38].

2) *Complex-Valued Rotationally Invariant Noise*: If we make the assumption that the (complex) noise is rotationally invariant, i.e., $p(w = \rho e^{j\theta}) = p(|w| = \rho)/2\pi$ for all $\theta \in [0, 2\pi]$, then we have that $E\{w^n\} = E\{\rho^n\} E\{e^{jn\theta}\} = 0$, for $n = 1, 2, \dots$. Using this assumption the cost function

reduces to

$$\begin{aligned} J_{\text{CM}}|_{\text{r.i. noise}} &= \kappa_s \sigma_s^4 \sum_{i=0}^{P-1} |h_i|^4 + 2\sigma_s^4 \sum_{i=0}^{P-1} \sum_{m=0, m \neq i}^{P-1} \\ &\cdot |h_i|^2 |h_m|^2 + |E\{s^2\}|^2 \sum_{i=0}^{P-1} \sum_{j=0, j \neq i}^{P-1} \\ &\cdot h_i^2 (h_j^*)^2 + \kappa_w \sigma_w^4 \sum_{i=0}^{2N-1} |f_i|^4 + 2\sigma_w^4 \sum_{i=0}^{2N-1} \\ &\cdot \sum_{m=0, m \neq i}^{2N-1} |f_i|^2 |f_m|^2 + 4\sigma_s^2 \sigma_w^2 \|\mathbf{h}\|_2^2 \|\mathbf{f}\|_2^2 \\ &- 2\sigma_s^2 \kappa_s (\sigma_s^2 \|\mathbf{h}\|_2^2 + \sigma_w^2 \|\mathbf{f}\|_2^2) + \sigma_s^4 \kappa_s^2 \end{aligned} \quad (64)$$

For the remaining derivations we will make the assumption of rotationally invariant noise.

3) *PSK Source*: For PSK symbols, $s_n \in \{e^{j2\pi m/2^M}\}$ and $m \in \{0, 1, \dots, 2^M - 1\}$ (where $j = \sqrt{-1}$), we note that $\sigma_s^2 = E\{|s|^4\} = \kappa_s = \sigma_s^4 = 1$. Thus, (64) simplifies to

$$\begin{aligned} J_{\text{CM}}|_{\text{PSK}} &= \sum_{i=0}^{P-1} |h_i|^4 + 2 \sum_{i=0}^{P-1} \sum_{m=0, m \neq i}^{P-1} |h_i|^2 |h_m|^2 \\ &+ |E\{s^2\}|^2 \sum_{i=0}^{P-1} \sum_{j=0, j \neq i}^{P-1} h_i^2 (h_j^*)^2 + \kappa_w \sigma_w^4 \\ &\cdot \sum_{i=0}^{2N-1} |f_i|^4 + 2\sigma_w^4 \sum_{i=0}^{2N-1} \sum_{m=0, m \neq i}^{2N-1} |f_i|^2 |f_m|^2 \\ &+ 4\sigma_w^2 \|\mathbf{h}\|_2^2 \|\mathbf{f}\|_2^2 - 2(\|\mathbf{h}\|_2^2 + \sigma_w^2 \|\mathbf{f}\|_2^2) + 1. \end{aligned} \quad (65)$$

4) *QAM Source*: For 90° rotationally invariant QAM (i.e., for every member q_m in the QAM alphabet, $\{jq_m, -q_m, -jq_m\}$ are equally likely members of the alphabet), we have that $E\{s^2\} = 0$ and (64) reduces to

$$\begin{aligned} J_{\text{CM}}|_{\text{QAM}} &= \kappa_s \sigma_s^4 \sum_{i=0}^{P-1} |h_i|^4 + 2\sigma_s^4 \sum_{i=0}^{P-1} \sum_{m=0, m \neq i}^{P-1} |h_i|^2 \\ &\cdot |h_m|^2 + \kappa_w \sigma_w^4 \sum_{i=0}^{2N-1} |f_i|^4 + 2\sigma_w^4 \sum_{i=0}^{2N-1} \\ &\cdot \sum_{m=0, m \neq i}^{2N-1} |f_i|^2 |f_m|^2 + 4\sigma_s^2 \sigma_w^2 \|\mathbf{h}\|_2^2 \|\mathbf{f}\|_2^2 \\ &- 2\sigma_s^2 \kappa_s (\sigma_s^2 \|\mathbf{h}\|_2^2 + \sigma_w^2 \|\mathbf{f}\|_2^2) + \sigma_s^4 \kappa_s^2. \end{aligned} \quad (66)$$

APPENDIX III

CM-MINIMIZING EQUALIZATION LITERATURE

- 1) R. A. Axford, Jr., L. B. Milstein, and J. R. Zeidler, "On the misconvergence of CMA blind equalizers in the reception of PN sequences," in *Proc. IEEE Military Commun. Conf.*, Fort Monmouth, NJ, Oct. 2-5, 1994, pp. 281-286. **Categories B, S, R.**
- 2) —, "Refined techniques for blind equalization of phase-shift keyed (PSK) and quadrature amplitude modulated QAM

- digital communications signals," Ph.D. dissertation, Univ. of California, San Diego, 1995. **Categories B, S, E.**
- 3) R. A. Axford, Jr., L. B. Milstein, and J. R. Zeidler, "The effects on PN sequences on the misconvergence of the constant modulus algorithm," *IEEE Trans. Signal Processing*, vol. 46, pp. 519–523, Feb. 1998. **Categories B, S, R.**
 - 4) 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. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Albuquerque, NM, Apr. 1990, pp. 1675–1678. **Categories B, N.**
 - 5) C. K. Chan and J. J. Shynk, "Stationary points of the constant modulus algorithm for real Gaussian signals," *IEEE Trans. Acoustics, Speech, Signal Processing*, vol. 38, pp. 2176–2181, Dec. 1990. **Categories B, S, T.**
 - 6) W. Chung and J. P. LeBlanc, "The local minima of fractionally spaced CMA blind equalizer cost function in the presence of channel noise," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Seattle, WA, May 1998, pp. 3345–3348. **Categories F, T, N.**
 - 7) R. Cusani and A. Laurenti, "Evaluation of the constant modulus algorithm in blind equalization of three ray multipath fading channels," *European Trans. Telecommun. Related Tech.*, vol. 6, no. 2, pp. 187–90, Feb.–Apr. 1995.
 - 8) —, "Convergence analysis of the CMA blind equalizer," *IEEE Trans. Commun.*, vol. 43, pp. 1304–1307, Feb.–Apr. 1995. **Category B.**
 - 9) Z. Ding, C. R. Johnson, Jr., R. A. Kennedy, and B. D. O. Anderson, "On the ill-convergence of Godard blind equalizers in data communication systems," in *Proc. Conf. Inform. Science and Systems*, Baltimore, MD, Mar. 1989, pp. 538–543. **Categories B, L, T, I.**
 - 10) Z. Ding, C. R. Johnson, Jr. and, R. A. Kennedy, "Admissibility in blind adaptive equalizers," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Albuquerque, NM, Apr. 1990, pp. 1707–1710. **Category B.**
 - 11) Z. Ding, "Application aspects of blind adaptive equalizers in QAM data communications," Ph.D. dissertation, Cornell University, Ithaca, NY, 1990. **Categories B, L, T, I.**
 - 12) Z. Ding, C. R. Johnson, Jr., and R. A. Kennedy, "Local convergence of 'globally convergent' blind adaptive equalization algorithms," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Toronto, Ont., Canada, May 14–17, 1991, pp. 1533–1536. **Categories B, L, T.**
 - 13) Z. Ding and C. R. Johnson, Jr., "Existing gap between theory and application of blind equalization," in *Proc. Int. Soc. Optical Engineers*, San Diego, CA, July 1991, pp. 154–165. **Categories B, S, T.**
 - 14) Z. Ding, R. A. Kennedy, B. D. O. Anderson, and C. R. Johnson, Jr., "Ill-convergence of Godard blind equalizers in data communication systems," *IEEE Trans. Commun.*, vol. 39, pp. 1313–1327, Sept. 1991. **Categories B, L, T.**
 - 15) Z. Ding, C. R. Johnson, Jr., 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. **Categories B, L, T.**
 - 16) Z. Ding, and R. A. Kennedy, "On the whereabouts of local minima for blind adaptive equalizers," *IEEE Trans. Circuits Syst. II*, vol. 39, pp. 119–123, Feb. 1992. **Categories B, L, T.**
 - 17) Z. Ding and C. R. Johnson, Jr., "On the nonvanishing stability of undesirable equilibria for FIR Godard blind equalizers," *IEEE Trans. Signal Processing*, vol. 41, pp. 1940–1944, May 1993. **Categories B, L, T.**
 - 18) Z. Ding, C. R. Johnson, Jr., and R. A. Kennedy, "Global convergence issues with linear blind adaptive equalizers," in *Blind Deconvolution*. Englewood Cliffs, NJ: Prentice-Hall, 1994, pp. 60–120. **Categories B, L, T.**
 - 19) Z. Ding, "On convergence analysis of fractionally spaced adaptive blind equalizers," *IEEE Trans. Signal Processing*, vol. 45, pp. 650–657, Mar. 1997. **Categories F, P, T.**
 - 20) T. J. Endres, B. D. O. Anderson, C. R. Johnson, Jr., and M. Green, "On the robustness of the fractionally spaced constant modulus criterion to channel order undermodeling: Part I," in *Proc. IEEE Signal Processing Workshop Signal Processing Advances Wireless Communications*, Paris, France, Apr. 16–18, 1997, pp. 37–40. **Categories F, L, T.**
 - 21) —, "On the robustness of the fractionally spaced constant modulus criterion to channel order undermodeling: Part II," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Munich, Germany, Apr. 20–24, 1997, pp. 3605–3608. **Categories F, L, T, E.**
 - 22) T. J. Endres, "Equalizing with fractionally spaced constant modulus and second-order-statistics blind receivers," Ph.D. dissertation, Cornell University, Ithaca, NY, 1997. **Categories F, L, T, E.**
 - 23) S. Evans and L. Tong, "Adaptive channel surfing re-initialization of the constant modulus algorithm," in *Proc. Asimolar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 1997. **Categories B, I.**
 - 24) I. Fijalkow, F. Lopez de Victoria, and C. R. Johnson, Jr., "Adaptive fractionally spaced blind equalization," in *Signal Processing Workshop*, Yosemite National Park, CA, Oct. 2–5, 1994, pp. 257–260. **Categories F, P, T.**
 - 25) I. Fijalkow, J. R. Treichler, and C. R. Johnson, Jr., "Fractionally spaced blind equalization: Loss of channel disparity," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Detroit, MI, May 9–12, 1995, pp. 1988–1991. **Categories F, D.**
 - 26) I. Fijalkow, A. Touzni, and C. R. Johnson, Jr., "Spatio-temporal equalizability under channel noise and loss of disparity," in *Proc. Colloque GRETSI Traitement du Signal et des Images*, Sophia Antipolis, France, Sept. 1995, pp. 293–296. **Categories F, N, D.**
 - 27) I. Fijalkow, A. Touzni, and J. R. Treichler, "Fractionally spaced equalization using CMA: Robustness to channel noise and lack of disparity," *IEEE Trans. Signal Processing*, vol. 45, pp. 56–66, Jan. 1997. **Categories F, N, D, T.**
 - 28) I. Fijalkow, C. Manlove, and C. R. Johnson, Jr., "Adaptive fractionally spaced blind CMA equalization: Excess MSE," *IEEE Trans. Signal Processing*, vol. 46, pp. 227–231, Jan. 1998. **Categories F, P, E.**
 - 29) G. J. Foschini, "Equalizing without altering or detecting data (digital radio systems)," *AT&T Tech. J.*, vol. 64, no. 8, pp. 1885–1911, Oct. 1985. **Categories B, P, T, R.**
 - 30) M. R. Frater, R. R. Bitmead, and C. R. Johnson, Jr., "Escape from stable equilibria in blind adaptive equalizers," in *Proc. IEEE Conf. Decision Control*, Tucson, AZ, Dec. 1992, pp. 1756–1761. **Categories B, T, R.**
 - 31) M. R. Frater and C. R. Johnson, Jr., "Local minima escape transients of CMA," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Adelaide, SA, Australia, Apr. 19–22, 1994, pp. 37–40. **Categories B, T, R.**
 - 32) M. R. Frater, R. R. Bitmead, and C. R. Johnson, Jr., "Local minima escape transients by stochastic gradient descent algorithms in blind adaptive equalizers," *Automatica*, vol. 31, no. 4, pp. 637–641, Apr. 1995. **Categories B, T, R.**
 - 33) D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. 28, pp. 1867–1875, Nov. 1980. **Categories B, T, I, P.**
 - 34) Z. Gu and W. A. Sethares, "A geometrical view of blind equalization," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Minneapolis, MN, Apr. 27–30, 1993, vol. 3, pp. 551–554. **Categories B, T.**
 - 35) K. Hilal and P. Duhamel, "A general form for recursive adaptive algorithms leading to an exact recursive CMA," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, San Francisco, CA, Mar. 1992, pp. 17–20. **Categories B, R.**
 - 36) —, "A convergence study of the constant modulus algorithm leading to a normalized-CMA and a block-normalized-CMA," in *Proc. European Signal Processing Conf.*, Brussels, Belgium, Aug. 24–27, 1992, vol. 1, pp. 135–138. **Categories B, R.**
 - 37) —, "Stability and convergence analysis of the constant modulus algorithm. Comments on finite equalization schemes and the stability of the normalized CMA," in *Proc. European Signal Processing Conf.*, Edinburgh, UK, Sept. 13–16, 1994, pp. 724–727.
 - 38) H. Jamali and S. L. Wood, "Error surface analysis for the complex constant modulus adaptive algorithm," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 1990, pp. 248–252. **Categories B, T.**
 - 39) H. Jamali, S. L. Wood, and R. Cristi, "Experimental validation of the Kronecker product Godard blind adaptive algorithms," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Oct. 26–28, 1992, vol. 1, pp. 1–5. **Categories B, T, N, L.**

- 40) C. R. Johnson, Jr., S. Dasgupta, and W. A. Sethares, "Averaging analysis of local stability of a real constant modulus algorithm adaptive filter," *IEEE Trans. Acoustics, Speech, Signal Processing*, vol. 36, pp. 900–910, June 1988. **Category B**.
- 41) C. R. Johnson, Jr., "Admissibility in blind adaptive channel equalization," *IEEE Contr. Syst. Mag.*, vol. 11, pp. 3–15, Jan. 1991. **Categories B, T, I**.
- 42) C. R. Johnson, Jr., P. C. E. Bennet, J. P. LeBlanc, and V. Krishnamurthy, "Blind adaptive equalizer average stationary point stability analysis with admissibility consequences," in *Proc. Conf. Information Science and Systems*, Princeton, NJ, Mar. 1992, pp. 727–732. **Categories B, S, L, I**.
- 43) C. R. Johnson, Jr. and J. P. LeBlanc, "Towards operational guidelines for memoryless-error-function-style blind equalizers," in *Proc. COST 229 Workshop Adaptive Algorithms Commun.*, Bordeaux-Technopolis, France, Sept. 1992, pp. 5–17. **Categories B, S**.
- 44) C. R. Johnson, Jr., J. P. LeBlanc, and V. Krishnamurthy, "Godard blind equalizer misbehavior with correlated sources," *J. Marocain d'Automat., d'Inform., Signal.*, vol. 2, pp. 1–39, June 1993. **Categories B, S**.
- 45) C. R. Johnson, Jr. and B. D. O. Anderson, "Godard blind equalizer error surface characteristics: White, zero-mean, binary case," *Int. J. Adaptive Contr. Signal Processing*, vol. 9, pp. 301–324, July–Aug. 1995. **Categories B, T, L**.
- 46) O. W. Kwon, C. K. Un, and J. C. Lee, "Performance of constant modulus adaptive digital filters for interference cancellation," *Signal Processing*, vol. 26, pp. 185–196, Feb. 1992. **Categories B, R, N**.
- 47) S. Lambotharan, J. Chambers, and C. R. Johnson, Jr., "Attraction of saddles and slow convergence in CMA adaptation," *Signal Processing*, vol. 59, pp. 335–340, June 1997. **Categories F, R, T, P**.
- 48) M. G. Larimore and J. R. Treichler, "Convergence behavior of the constant modulus algorithm," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Boston, MA, Apr. 14–16, 1983, pp. 13–16. **Categories B, R**.
- 49) —, "The capture properties of CMA-based interference cancellers," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 5–7, 1984, pp. 49–52. **Categories B, S, I, N**.
- 50) —, "Noise capture properties of constant modulus algorithm," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Tampa, FL, Mar. 26–29, 1985, pp. 1165–1168. **Categories B, N, S**.
- 51) J. P. LeBlanc, K. Dogancay, R. A. Kennedy, and C. R. Johnson, Jr., "Effects of input data correlation on the convergence of blind adaptive equalizers," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Adelaide, SA, Australia, Apr. 19–22, 1994, pp. 313–316. **Categories B, R, S**.
- 52) J. P. LeBlanc, I. Fijalkow, B. Huber, and C. R. Johnson, Jr., "Fractionally spaced CMA equalizers under periodic and correlated inputs," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Detroit, MI, May 9–12, 1995, pp. 1041–1044. **Categories F, S, T**.
- 53) J. P. LeBlanc, "Effects of source distributions and correlation on fractionally spaced blind constant modulus algorithm equalizers," Ph.D. dissertation, Cornell University, Ithaca, NY, 1995. **Categories F, S, T, R**.
- 54) J. P. LeBlanc and S. W. McLaughlin, "Non-equiprobable constellation shaping and blind constant modulus algorithm equalization," in *Proc. Conf. Information Science and Systems*, Princeton, NJ, Mar. 1996, pp. 901–903. **Categories F, S**.
- 55) J. P. LeBlanc, I. Fijalkow, and C. R. Johnson, Jr., "Fractionally spaced constant modulus algorithm blind equalizer error surface characterization: Effects of source distributions," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Atlanta, GA, May 7–9, 1996, pp. 2944–2947. **Categories F, S, T, R**.
- 56) J. P. LeBlanc and C. R. Johnson, Jr., "Global CMA error surface characteristics, source statistic effects: Polytopes and manifolds," in *Proc. Int. Conf. Digital Signal Processing*, Santorini, Greece, July 2–4, 1997, pp. 131–134. **Categories F, S, T, P**.
- 57) J. P. LeBlanc, I. Fijalkow, and C. Richard Johnson, Jr., "CMA fractionally spaced equalizers: Stationary points and stability under IID and temporally correlated sources," *Int. J. Adaptive Contr. and Signal Processing*, vol. 12, no. 2, pp. 135–155, March 1998. **Categories F, S, T**.
- 58) Y. Li and Z. Ding, "Global convergence of fractionally spaced Godard equalizers," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Oct. 31–Nov. 2, 1994, vol. 1, pp. 617–621. **Categories F, T, P**.
- 59) —, "Convergence analysis of finite length blind adaptive equalizers," *IEEE Trans. Signal Processing*, vol. 43, pp. 2120–2129, Sept. 1995. **Categories B, T, L, I**.
- 60) —, "Global convergence of fractionally spaced Godard (CMA) adaptive equalizers," *IEEE Trans. Signal Processing*, vol. 44, pp. 818–826, Apr. 1996. **Categories F, T, P**.
- 61) Y. Li, K. J. R. Liu, and Z. Ding, "Length and cost dependent local minima of unconstrained blind channel equalizers," *IEEE Trans. Signal Processing*, vol. 44, pp. 2726–2735, Nov. 1996. **Categories B, T, L**.
- 62) Y. Li and K. J. R. Liu, "Static and dynamic convergence behavior of adaptive blind equalizers," *IEEE Trans. Signal Processing*, vol. 22, pp. 2736–2745, Nov. 1996. **Categories B, L, E**.
- 63) F. Lopez de Victoria, A. Bosser, I. Fijalkow, C. R. Johnson, Jr., and J. R. Treichler, "Observed (mis)behavior of CMA with periodic sources: Assessment and guidelines," in *Proc. IEEE Signal Processing Workshop*, Yosemite National Park, CA, Oct. 2–5, 1994, pp. 261–264. **Categories B, S**.
- 64) W. E. Meyer and J. P. LeBlanc, "Blind adaptive fractionally spaced CMA in the presence of channel noise," in *Proc. Conf. Information Science and Systems*, Princeton, NJ, Mar. 1996, pp. 373–374. **Categories F, N**.
- 65) C. Papadias, "On the existence of undesired global minima of Godard equalizers," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Munich, Germany, Apr. 20–24, 1997, pp. 3937–3940. **Categories F, T, S, P**.
- 66) J. J. Shynk and C. K. Chan, "A comparative analysis of the stationary points of the constant modulus algorithm based on Gaussian assumptions," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Albuquerque, NM, Apr. 1990, pp. 1249–1252. **Categories B, S, T**.
- 67) —, "Error surfaces of the constant modulus algorithm," in *Int. Symposium Circuits and Systems*, New Orleans, LA, May 1990, pp. 1335–1338. **Categories B, T, S**.
- 68) —, "Performance surfaces of the constant modulus algorithm based on a conditional Gaussian model," *IEEE Trans. Signal Processing*, vol. 41, pp. 1965–1969, May 1993. **Categories B, S, T**.
- 69) J. O. Smith and B. Freidlander, "Global convergence of the constant modulus algorithm," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Tampa, FL, Mar. 26–29, 1985, pp. 1161–1164. **Categories B, P, T**.
- 70) R. Swaminathan and J. K. Tugnait, "On improving the convergence of constant modulus algorithm adaptive filters," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Minneapolis, MN, Apr. 27–30, 1993, pp. 340–343. **Categories B, I**.
- 71) L. Tong and H. Zeng, "Channel surfing re-initialization for the constant modulus algorithm," *IEEE Signal Processing Lett.*, vol. 4, pp. 85–87, Mar. 1997. **Categories F, I**.
- 72) A. Touzni, I. Fijalkow, and J. R. Treichler, "Fractionally spaced CMA under channel noise," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Atlanta, GA, May 7–9, 1996, pp. 2674–2677. **Categories F, N, T**.
- 73) —, "Robustness of fractionally spaced equalization by CMA to lack of channel disparity," in *Proc. IEEE Signal Processing Workshop Statistical Signal Array Processing*, Corfu, Greece, June 24–26, 1996, pp. 144–147. **Categories F, D, T**.
- 74) A. Touzni and I. Fijalkow, "Does fractionally spaced CMA converge faster than LMS?" in *Proc. European Signal Processing Conf.*, Trieste, Italy, Sept. 1996, pp. 1227–1230. **Categories F, R, T, P**.
- 75) —, "Channel robust blind fractionally spaced equalization," in *IEEE Signal Processing Workshop Signal Processing Advances Wireless Commun.*, Paris, France, Apr. 1997, pp. 33–36. **Categories F, D, I**.
- 76) —, "Robustness of blind fractionally spaced identification/equalization to loss of channel disparity," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Munich, Germany, Apr. 1997, pp. 3937–3940. **Categories F, D, T**.
- 77) J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoustics, Speech, Signal Processing*, vol. ASSP-31, pp. 459–472, Apr. 1983. **Categories B, T, I, P**.

- 78) J. R. Treichler and M. G. Larimore, "Convergence rates for the constant modulus algorithm with sinusoidal inputs," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Tampa, FL, Mar. 26–29, 1985, pp. 1157–1160. **Categories B, R, S.**
- 79) —, "The tone capture properties of CMA-based interference suppressors," *IEEE Trans. Acoustics, Speech, Signal Processing*, vol. ASSP-33, pp. 946–958, Aug. 1985. **Categories B, S.**
- 80) —, "New processing techniques based on the constant modulus adaptive algorithm," *IEEE Trans. Acoustics, Speech, Signal Processing*, vol. ASSP-33, pp. 420–431, Apr. 1985. **Categories B, F.**
- 81) J. R. Treichler, V. Wolff, and C. R. Johnson, Jr., "Observed misconvergence in the constant modulus adaptive algorithm," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 1991, pp. 663–667. **Category S.**
- 82) J. R. Treichler, L. Tong, I. Fijalkow, C. R. Johnson, Jr., and C. U. Berg, "On the current shape of FSE-CMA behavior theory," in *Proc. IEEE Signal Processing Workshop Signal Processing Advances Wireless Communications*, Paris, France, Apr. 1997, pp. 105–108. **Categories F, N, L.**
- 83) J. K. Tugnait, "A parallel multimodel: CMA/Godard adaptive filter bank approach to fractionally spaced blind adaptive equalization," in *Proc. IEEE Int. Conf. Communications*, New Orleans, LA, May 1–5, 1994, vol. 1, pp. 549–553. **Category F.**
- 84) —, "On fractionally spaced blind adaptive equalization under symbol timing offsets using Godard and related equalizers," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Detroit, MI, May 9–12, 1995, vol. 3, pp. 1976–1979. **Category F.**
- 85) —, "On fractionally spaced blind adaptive equalization under symbol timing offsets using Godard and related equalizers," *IEEE Trans. Signal Processing*, vol. 44, pp. 1817–1821, July 1996. **Categories B, F.**
- 86) S. Vembu, S. Verdu, R. A. Kennedy, and W. A. Sethares, "Convex cost functions in blind equalization," *IEEE Trans. Signal Processing*, vol. 42, pp. 1952–1960, Aug. 1994. **Category B.**
- 87) M. Wu and F. Cornett, "Discrete-time and continuous-time constant modulus algorithm analysis," in *Proc. Southeastern Symp. Systems Theory*, Starkville, MS, Mar. 12–14, 1995, pp. 504–508. **Categories B, R, E.**
- 88) H. H. Zeng and L. Tong, "On the performance of CMA in the presence of noise—Some new results on blind channel estimation: Performance and algorithms," in *Proc. Conf. Information Science and Systems*, Princeton, NJ, Mar. 1996, pp. 890–894. **Categories F, N, T.**
- 89) H. H. Zeng, L. Tong, and C. R. Johnson, Jr., "Behavior of fractionally spaced constant modulus algorithm: Mean square error, robustness and local minima," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 1996, pp. 305–309. **Categories F, N, T, D, L.**
- 90) H. H. Zeng and L. Tong, "The MSE performance of constant modulus receivers," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Munich, Germany, Apr. 1997, pp. 3577–3580. **Categories B, F, N, T.**
- 91) H. H. Zeng, L. Tong, and C. R. Johnson, Jr., "Relationships between the constant modulus and Wiener receivers," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1523–1538, July 1998. **Categories F, N, T, D, L.**
- 92) S. Zeng, H. H. Zeng, and L. Tong, "Blind equalization using CMA: Performance analysis and a new algorithm," in *Proc. IEEE Int. Conf. Communications*, Dallas, TX, June 1996, pp. 847–851. **Categories F, N, T.**
- 93) E. Zervas, J. G. Proakis, and V. Eyuboglu, "Effect of constellation shaping on blind equalization," in *Proc. Int. Soc. Optical Engineers*, San Diego, CA, pp. 1393–1397, 1991. **Categories B, S.**

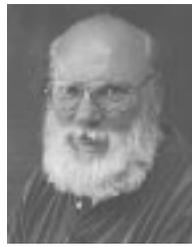
ACKNOWLEDGMENT

The authors would like to thank the following people for their comments on earlier versions of this paper: C. Papadias, J. Treichler, I. Fijalkow, A. Sayed, D. Jones, A. M. Baksho, W. Chung, J. Balakrishnan, M. A. Poubelle, C. P. Hankey, and A. F. Bob.

REFERENCES

- [1] J. W. M. Bergmans and A. J. E. M. Janssen, "Robust data equalization, fractional tap spacing and the Zak transform," *Philips J. Res.*, vol. 42, no. 4, pp. 351–398, 1987.
- [2] Z. Ding, "Characteristics of band-limited channels unidentifiable from second-order cyclostationary statistics," *IEEE Signal Processing Lett.*, vol. 3, pp. 150–152, May 1996.
- [3] G. D. Forney, L. Brown, M. V. Eyuboglu, and J. L. Moran III, "The V.34 high-speed modem standard," *IEEE Commun. Mag.*, vol. 34, pp. 28–33, Dec. 1996.
- [4] R. D. Gitlin, "Fractionally spaced equalization: An improved digital transversal equalizer," *Bell Syst. Tech. J.*, vol. 60, pp. 275–296, Feb. 1981.
- [5] R. D. Gitlin, J. F. Hayes, and S. B. Weinstein, *Data Communications Principles*. New York: Plenum Press, 1992.
- [6] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [7] C. R. Johnson, Jr. et al., "On fractionally spaced equalizer design for digital microwave radio channels," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, Oct. 1995, pp. 290–294.
- [8] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1980.
- [9] M. G. Larimore and M. J. Goodman, "Implementation of the constant modulus algorithm at RF bandwidths," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, Nov. 1985, pp. 626–630.
- [10] E. A. Lee and D. G. Messerschmitt, *Digital Communication*, 2nd ed. Boston, MA: Kluwer, 1994.
- [11] R. W. Lucky, "Techniques for adaptive equalization of digital communication systems," *Bell Syst. Techn. J.*, vol. 45, pp. 255–286, Feb. 1966.
- [12] —, "Signal filtering with the transversal equalizer," in *Proc. Allerton Conf. Circuits Syst. Theory*, Allerton, IL, Oct. 1969, p. 792.
- [13] D. G. Luenberger, *Optimization by Vector Space Methods*, 2nd ed. New York: Wiley, 1990.
- [14] O. Macchi and L. Guidoux, "A new equalizer: The double-sampling equalizer," *Annales des Telecommun.*, vol. 30, nos. 9/10, pp. 331–338, Sept. 1975.
- [15] E. Moulines, P. Duhamel, J. Cardoso, and S. Mayrargue, "Subspace methods for blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 516–525, Feb. 1995.
- [16] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.
- [17] S. U. H. Qureshi, "Adjustment of the position of the reference tap of an adaptive equalizer," *IEEE Trans. Commun.*, pp. 1046–1052, Sept. 1973.
- [18] —, "Adaptive equalization," *Proc. IEEE*, vol. 73, pp. 1349–1387, Sept. 1985.
- [19] Y. Sato, "A method of self-recovering equalization for multi-level amplitude-modulations systems," *IEEE Trans. Commun.*, pp. 679–682, June 1975.
- [20] G. Strang, *Linear Algebra and its Applications*, 3rd ed. Fort Worth, TX: Harcourt Brace Jovanovich, 1988.
- [21] L. Tong, "A fractionally spaced adaptive blind equalizer," in *Proc. Conf. Inform. Sci. Syst.*, Princeton, NJ, Mar. 1992, pp. 711–716.
- [22] L. Tong, G. Xu, B. Hassibi, and T. Kailath, "Blind identification and equalization based on second-order statistics: A frequency-domain approach," *IEEE Trans. Inform. Theory*, vol. 41, pp. 329–334, Jan. 1995.
- [23] J. R. Treichler, I. Fijalkow, and C. R. Johnson, Jr., "Fractionally spaced equalizers: How long should they really be?," *IEEE Signal Processing Mag.*, vol. 13, pp. 65–81, May 1996.
- [24] J. R. Treichler, M. G. Larimore, and J. C. Harp, "Practical blind demodulators for high-order QAM signals," this issue, pp. 1907–1926.
- [25] J. K. Tugnait, "On blind identifiability of multipath channels using fractional sampling and 2nd-order cyclostationary statistics," *IEEE Trans. Inform. Theory*, vol. 41, pp. 308–311, Jan. 1995.
- [26] G. Ungerboeck, "Fractional tap spacing equalizer and consequences for clock recovery in data modems," *IEEE Trans. Commun.*, vol. 24, no. 8, pp. 856–864, Aug. 1976.

- [27] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Engelwood Cliffs, NJ: Prentice-Hall, 1985.
- [28] V. G. Wolff, R. P. Gooch, and J. R. Treichler, "Specification and development of an equalizer-demodulator for wideband digital microwave radio signals," in *Proc. IEEE Military Commun. Conf.*, San Diego, CA, Oct. 1988, pp. 461–467.
- [29] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. 28, pp. 1867–1875, Nov. 1980.
- [30] J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoustics, Speech, Signal Processing*, vol. 31, pp. 459–72, Apr. 1983.
- [31] G. J. Foschini, "Equalizing without altering or detecting data (digital radio systems)," *AT&T Tech. J.*, vol. 64, no. 8, pp. 1885–1911, Oct. 1985.
- [32] Y. Li and Z. Ding, "Global convergence of fractionally spaced Godard (CMA) adaptive equalizers," *IEEE Trans. Signal Processing*, vol. 44, no. 4, pp. 818–826, Apr. 1996.
- [33] I. Fijalkow, A. Touzni, and J. R. Treichler, "Fractionally spaced equalization using CMA: Robustness to channel noise and lack of disparity," *IEEE Trans. Signal Processing*, vol. 45, no. 1, pp. 56–66, Jan. 1997.
- [34] M. G. Larimore and J. R. Treichler, "Convergence behavior of the constant modulus algorithm," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Boston, MA, Apr. 14–16, 1983, pp. 13–16.
- [35] S. Lambotharan, J. Chambers, and C. R. Johnson, Jr., "Attraction of saddles and slow convergence in CMA adaptation," *Signal Processing*, vol. 59, no. 2, pp. 335–340, June 1997.
- [36] A. Touzni and I. Fijalkow, "Does fractionally spaced CMA converge faster than LMS?," in *Proc. European Signal Processing Conf.*, Trieste, Italy, Sept. 1996, pp. 1227–1230.
- [37] I. Fijalkow, C. Manlove, and C. R. Johnson, Jr., "Adaptive fractionally spaced blind CMA equalization: Excess MSE," *IEEE Trans. Signal Processing*, vol. 46, pp. 227–231, Jan. 1998.
- [38] C. R. Johnson, Jr. and B. D. O. Anderson, "Godard blind equalizer error surface characteristics: White, zero-mean, binary case," *Int. J. Adaptive Contr. Signal Processing*, vol. 9, pp. 301–324, July/Aug. 1995.
- [39] J. P. LeBlanc, I. Fijalkow, and C. Richard Johnson, Jr., "CMA fractionally spaced equalizers: Stationary points and stability under IID and temporally correlated sources," *International Journal of Adaptive Contr. Signal Processing*, 1998, to appear.
- [40] H. H. Zeng, L. Tong, and C. R. Johnson, Jr., "Relationships between the constant modulus and Wiener receivers," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1523–1538, July 1998.
- [41] W. Chung and J. P. LeBlanc, "The local minima of fractionally spaced CMA blind equalizer cost function in the presence of channel noise," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Seattle, WA, May 1998, pp. 3345–3348.
- [42] Y. Li and K. J. R. Liu, "Static and dynamic convergence behavior of adaptive blind equalizers," *IEEE Trans. Signal Processing*, vol. 44, pp. 2736–2745, Nov. 1996.
- [43] T. J. Endres, B. D. O. Anderson, C. R. Johnson, Jr., and M. Green, "On the robustness of the fractionally spaced constant modulus criterion to channel order undermodeling: Part I," in *Proc. IEEE Signal Processing Workshop on Signal Processing Advances Wireless Commun.*, Paris, France, Apr. 16–18, 1997, pp. 37–40.
- [44] ———, "On the robustness of the fractionally spaced constant modulus criterion to channel order undermodeling: Part II," in *Proc. IEEE Int. Conf. Acoustics, Speech Signal Proc.*, Munich, Germany, Apr. 20–24, 1997, pp. 3605–3608.
- [45] J. R. Treichler, V. Wolff, and C. R. Johnson, Jr., "Observed misconvergence in the constant modulus adaptive algorithm," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 1991, pp. 663–667.
- [46] R. A. Axford, Jr., L. B. Milstein, and J. R. Zeidler, "The effects on PN sequences on the misconvergence of the constant modulus algorithm," *IEEE Trans. Signal Processing*, vol. 46, pp. 519–523, Feb. 1998.
- [47] J. P. LeBlanc, "Effects of source distributions and correlation on fractionally spaced blind constant modulus algorithm equalizers," Ph.D. dissertation, Cornell Univ., Ithaca, NY, 1995.



C. Richard Johnson, Jr. (Fellow, IEEE) 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 Engineering and a member of the Graduate Field of Applied Mathematics at Cornell University, Ithaca, NY. His current research interest is in adaptive parameter estimation theory, which is useful in applications of digital signal processing to telecommunication systems. His principal focus in the 1990's has been blind linear equalization for intersymbol interference removal from received QAM sources.



Philip Schniter 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 at Urbana-Champaign in 1992 and 1993. Since 1996 he has been working toward the Ph.D. degree in electrical engineering at Cornell University, Ithaca, NY

From 1993 to 1996 he was employed by Tektronix, Inc., Beaverton, OR, as a Systems Engineer. While there he worked on signal processing aspects of video and communications instrumentation design, including algorithms, software, and hardware architectures. His research interests are in signal processing, communications, and control, and include blind adaptive equalization.

Mr. Schniter is a recipient of the Schlumberger Fellowship.



Thomas J. Endres (Member, IEEE) received the B.S. degree from Cornell University, Ithaca, NY, in 1990, the M.S. degree from the University of Southern California, Los Angeles, 1994, and the Ph.D. degree from Cornell University in 1997, all in electrical engineering.

From 1990 to 1994, he was employed by Hughes Space and Communications, Los Angeles, CA. Since 1997 he has been a Founding Member of Sarnoff Digital Communications, Newtown, PA. His current research interests include blind equalization and adaptive systems.



James D. Behm received the M.S. degree in mathematics in 1975 from the University of Arizona, Tucson, and the M.S. degree in computer science in 1986 from Johns Hopkins Whiting School of Engineering, Laurel, MD.

Since 1975, he has been a Mathematician with the Department of Defense, Ft. Meade, MD. In Spring, 1997 he was a Visiting Researcher at Cornell University, Ithaca, NY. His current research interests include signal processing and communications.



Donald R. Brown received the B.S. and M.S. degrees in electrical engineering from the University of Connecticut, Storrs, in 1992 and 1996, respectively. He is currently a graduate student at Cornell University, Ithaca, NY.

From 1992 to 1997 he worked for the General Electric Company, Plainville, CT, as a Development Engineer. His research interests include adaptive signal processing and communications.



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

His thesis work is a result of collaborative research with the United Technologies Research Center on identification and control of nonlinear combustion chamber dynamics. His current research interests include the analysis of signal processing, communications, and control related applications from a nonlinear dynamical systems perspective.

Mr. Casas is a recipient of the National Science Foundation Minorities Fellowship.