An Analysis of Constant Modulus Receivers

Hanks H. Zeng, *Member, IEEE,* Lang Tong, *Member, IEEE,* and C. Richard Johnson, Jr., *Fellow, IEEE*

***Abstract*— This paper investigates connections between (nonblind) Wiener receivers and blind receivers designed by minimizing the constant modulus (CM) cost. Applicable to both T-spaced and fractionally spaced FIR equalization, the main results include**

1) **a test for the existence of CM local minima near Wiener receivers;**
2) **an analytical description of CM receivers in the neighborhood of Wiener receivers;**
3) **mean square error (MSE) bounds for CM receivers.**

When the channel matrix is invertible, we also show that the CM receiver is approximately colinear with the Wiener receiver and provide a quantitative measure of the size of neighborhoods that contain the CM receivers and the accuracy of the MSE bounds.

***Index Terms*—Blind equalization, CMA, MSE.**

I. Introduction

BLIND equalization of intersymbol interference (ISI) in communication channels and blind separation of multiple users are promising signal processing techniques in certain communication system designs. One of the earliest blind receiver designs, and perhaps the most widely used in practice, is the Godard or the constant modulus algorithm (CMA) [8], [11], [18]. In his original paper, Godard observed in simulation that receivers designed by minimizing the constant modulus cost have similar MSE performance to the nonblind Wiener receivers. This striking observation provides strong support for using CM blind receivers because they not only do not require the cooperation of the transmitter but also achieve near optimal performance (in the sense of minimizing mean square error of the estimation). Similar observations was also made by Treichler and Agee [18].

Most early analyses of CMA exclude additive channel noise. It has been shown that CM receiver converges globally to the channel inverse when the channel matrix is full column rank, which includes doubly infinite T-spaced equalizers [7] and finite-length fractionally spaced equalizers [14]. In such cases, the channel inverse is the Wiener receiver when channel noise is not present. For finite-length T-spaced CMA equalization, however, the existence of local minima has been shown by Ding *et al.* [1]–[4], [13] and has been summarized by Li *et al.* [15].

When noise cannot be ignored, analysis based on small noise perturbation has been obtained in several ways [5], [6], [15], [17]. Although this perturbation analysis does not quantify specific conditions under which the analysis is valid, it has been observed in simulation examples that the near optimal performance of CMA holds well for a wide range of signal-to-noise ratios. The first exact analysis that establishes the connection between CM and Wiener receivers appeared in was obtained recently [19], [21] for the special case that the channel matrix has full column rank. The application of this result is, unfortunately, limited because the full rank condition, satisfied in beam forming and certain fractionally spaced equalization problems, is not valid for T-spaced or fractionally spaced equalization with insufficient equalizer length.

The main contribution of this paper is the development of a systematic procedure for the analysis of CM receivers. Unlike the perturbation analysis, our approach does not involve approximations. As a generalization to the geometrical approach presented in [19] and [21], our approach can be applied to cases when the channel matrix is singular. Such generalization enables us to treat both T-spaced and fractionally spaced equalization within the same theoretical framework. While the approach used in this paper is similar in spirit to that presented in [19] and [21], the generalization is nontrivial because certain subspace constraints must be imposed on the CM optimization. Further, the analysis presented in this paper can also be applied to arbitrary real sources. Only binary source was considered in [19], [21]. A comparison between the results obtained for the general case and that for channels with an invertible channel matrix provides interesting insight into how the rank condition affects the behavior of CM algorithms. The main results of the analysis include

1) a test for the existence of CM local minima near Wiener receivers;
2) an analytical description of CM receivers in the neighborhood of Wiener receivers;
3) mean square error (MSE) bounds for CM receivers.

As demonstrated in [20], the theory developed in this paper can be of value in addressing several design issues in blind equalization. For example, the analytical procedure presented in this paper allows us to analyze the effects of noise, signal constellation, equalizer length, channel diversity, local minima, and model mismatch.

The rest of the paper is organized as follows. Section II presents a general system model and the constant modulus receiver. Section III derives the MSE bound for constant modulus receivers. Finally, a conclusion is given in Section V, and all the proofs are relegated to the Appendix.

Manuscript received July 16, 1997; revised December 21, 1998. H. H. Zeng and L. Tong were supported in part by the National Science Foundation under Contract NCR-9321813 and by the Office of Naval Research under Contract N00014-96-1-0895. C. R. Johnson, Jr. was supported by NSF Grant MIP-9509011 and Applied Signal Technology, Inc. Part of the work in this paper was presented at the 30th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, November 1996. The associate editor coordinating the review of this paper and approving it for publication was Editor-in-Chief Prof. José M. F. Moura.

H. H. Zeng is with AT&T Laboratories—Research, Red Bank, NJ 07701-7033 USA.

L. Tong and C. R. Johnson, Jr. are with the School of Electrical Engineering, Cornell University, Ithaca, NY 14853 USA.

Publisher Item Identifier S 1053-587X(99)08317-8.

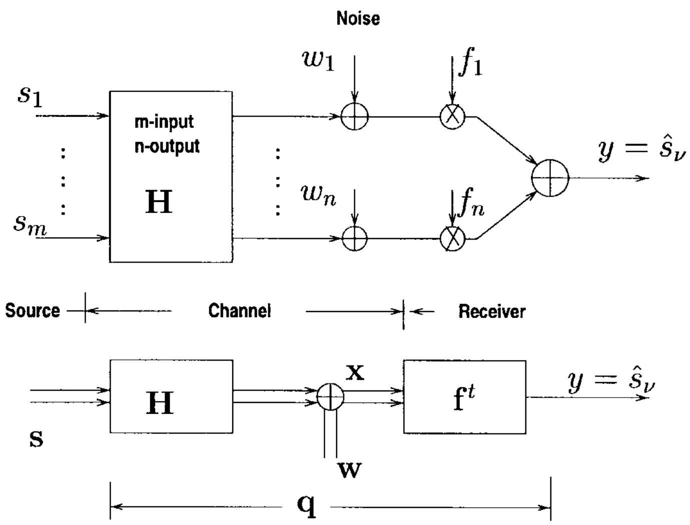


Fig. 1. Data model.

Most notation in this paper is standard with uppercase and lowercase bold letters denoting matrices and vectors, respectively. Special notations are listed as follows.

$(\cdot)^t$	Transpose.
$(\cdot)^\dagger$	Moore–Penrose inverse [12, p. 434].
$E\{\cdot\}$	Expectation operator.
$\|\|\boldsymbol{x}\|\|_p$	p-norm defined by $\sqrt[p]{\sum x_i^p}$.
$\|\|\boldsymbol{x}\|\|_{\boldsymbol{A}}$	2-norm defined by $\sqrt{\boldsymbol{x}^t\boldsymbol{A}\boldsymbol{x}}$.
$\boldsymbol{I}_n$	$n \times n$ identity matrix.
$\boldsymbol{e}_\nu$	Unit column vector with 1 at the νth entry and zero elsewhere.
$\mathcal{R}^n$	n-dimensional real vector space.
$\mathcal{R}^{n\times m}$	Set of all $n \times m$ real matrices.
$\mathcal{R}_{\boldsymbol{A}}$	Range of $\boldsymbol{A}\boldsymbol{A}^\dagger$ [12, p. 430].
$\mathcal{R}_{\boldsymbol{A}^\perp}$	Range of $\boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^\dagger$.
$\partial\mathcal{B}$	Boundary of set $\mathcal{B}$.

II. The Model

Constant modulus receivers can be applied to a broad class of applications such as blind equalization and beamforming. In this section, a general linear transmission model is given first followed by a generic CM receiver.

A. Data Model

We consider the estimation problem in the following linear model shown in Fig. 1. The system equation is given by

$$\boldsymbol{x} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{w}, \tag{1}$$

$$y = \boldsymbol{f}^t\boldsymbol{x} = \boldsymbol{q}^t\boldsymbol{s} + \boldsymbol{f}^t\boldsymbol{w} \tag{2}$$

where $\boldsymbol{s} = [s_1, \cdots, s_m]^t \in \mathcal{R}^m$ is a vector of the transmitted signal, $\boldsymbol{w} \in \mathcal{R}^n$ is the additive noise, $\boldsymbol{H} \in \mathcal{R}^{n\times m}$ is the unknown channel matrix, $\boldsymbol{x} \in \mathcal{R}^n$ is the received signal, $\boldsymbol{f} \in \mathcal{R}^n$ is the receiver parameter vector, y is the output of the receiver, and $\boldsymbol{q} \stackrel{\Delta}{=} \boldsymbol{H}^t\boldsymbol{f} \in \mathcal{R}^m$ is the combined channel-receiver response vector.

For equalization applications, vector $\boldsymbol{s}$ is composed of consecutive samples of the input, i.e., $\boldsymbol{s} = [s(k), s(k-1), \cdots, s(k-m+1)]^t$. The output y of the receiver is therefore an estimate of $s(k-\nu+1)$, which is the input with delay $\nu+1$. Note that the receiver delay can be specified in nonblind equalization problems. In contrast, in blind equalization algorithms such as CMA, the delay can only be controlled through algorithm initialization. Thus far, there is no systematic method of initialization that ensures convergence to the appropriate delay. The detailed derivation of $\boldsymbol{H}$ for both T-spaced and fractionally spaced equalization can be found in the Appendix.

We consider the rather general case when no restriction is imposed on the channel matrix $\boldsymbol{H}$. For the signals, we assume the following.

A0) All signals are real.
A1) $\boldsymbol{w}$ is zero mean Gaussian with covariance $\sigma^2\boldsymbol{I}$.
A2: Entries of $\boldsymbol{s}$ are independent random variables with $E\{s_\nu\} = 0$, $E\{s_\nu^2\} = 1$ and $E\{s_\nu^4\} = r$ $(0 < r < 3)$.

The restriction to the real case is not a fundamental one in the sense that the basic approach also applies to the complex case. However, most formulae and their interpretation may be some different in complex case. The transmitted signal s_ν is an arbitrary real signal, such as a symbol from binary phase-shift keying (BPSK) or multilevel pulse amplitude modulation (PAM) constellations. r is also referred to as the dispersion constant [8].

B. The Constant Modulus Receiver and CMA

In communication systems (see Fig. 1), the transmitted signal does not take on arbitrary values. For example, if the signal has a phase-shift-keying (PSK) modulation, s_ν is on the unit circle. Godard [8] and Treichler *et al.* [18] proposed the constant modulus (CM) criterion that minimizes the dispersion of the receiver output about the dispersion constant $r = E\{|s|^4\}/E\{|s|^2\}$

$$J_c(\boldsymbol{f}) \stackrel{\Delta}{=} E\left\{(|\underbrace{\boldsymbol{f}^t\boldsymbol{x}}_{y}|^2 - r)^2\right\}. \tag{3}$$

In our discussion, the local minima of $J_c(\boldsymbol{f})$ are referred to as constant modulus (CM) receivers.

In practical applications, a CM receiver is usually obtained from the stochastic gradient algorithm. The gradient of $J_c(\boldsymbol{f})$ is given by

$$\frac{\partial}{\partial \boldsymbol{f}} J_c(\boldsymbol{f}) = 2E\{(|\boldsymbol{f}^t\boldsymbol{x}|^2 - r)y\boldsymbol{x}\} \tag{4}$$

where $\boldsymbol{x}(k)$, $y(k)$, and $\boldsymbol{f}(k)$ be the channel output vector, the receiver output, and the receiver coefficient vector at time k, respectively. The constant modulus algorithm (CMA) is the stochastic gradient update of the receiver coefficients by removing the expectation operator in (4) and correcting $\boldsymbol{f}$ by a small amount in the opposite direction

$$\boldsymbol{f}(k+1) = \boldsymbol{f}(k) - \mu(|y(k)|^2 - r)y(k)\boldsymbol{x}(k). \tag{5}$$

According to the averaging analysis of [9], the mean CM cost function (3) describes the average performance of CMA in (5).

III. Mean Square Error of Constant Modulus Receivers

In this section, we develop a systematic procedure to locate the CM local minima and evaluate their mean square error (MSE) performance. Specifically, given $\boldsymbol{H}$, the signal-to-noise ratio, and signal constellation, we present an algorithm that enables us to test the existence of a CM receiver $\boldsymbol{f}_c$ in the neighborhood of the Wiener receiver to approximate its location to evaluate the upper bound of its MSE defined by

$$\mathcal{E}_c \stackrel{\Delta}{=} \mathrm{MSE}(y_c) = E\{|\boldsymbol{f}_c^t \boldsymbol{x} - s_\nu|^2\}. \tag{6}$$

To achieve this goal, we establish several key properties including the signal space property and the lower bound of the CM cost function in the neighborhood of the Wiener receiver.

A. Signal Space Property and Equivalent Cost Function

Under A1) and A2), the CM cost function has the following form, as shown in [10] and [21]:

$$\begin{aligned} J_c(\boldsymbol{f}) &\stackrel{\Delta}{=} E\{(|y|^2 - r)^2\} \\ &= 3\|\boldsymbol{f}\|_{\boldsymbol{R}}^4 - 2r\|\boldsymbol{f}\|_{\boldsymbol{R}}^2 - (3-r)\|\boldsymbol{H}^t\boldsymbol{f}\|_4^4 + r^2 \end{aligned} \tag{7}$$

where

$$\boldsymbol{R} \stackrel{\Delta}{=} E\{\boldsymbol{x}\boldsymbol{x}^t\} = \boldsymbol{H}\boldsymbol{H}^t + \sigma^2 \boldsymbol{I}_n. \tag{8}$$

CM receivers are defined as the local minima of the CM cost function $J_c(\boldsymbol{f})$.

One of the important properties of CM receivers is that they all must be in the "signal" subspace spanned by the columns of $\boldsymbol{H}$ (see also [21]), which implies that a CM receiver automatically has the matched filter front end.

Lemma 1: The output energy of any CMA receiver $\boldsymbol{f}_c$ satisfies

$$\frac{r}{3} < E\{\|y_c\|^2\} < 1. \tag{9}$$

Furthermore, all CMA local minima are identical to the local minima of the CMA cost function constrained in the signal subspace, i.e.,

$$\min J_c(\boldsymbol{f}),\ \boldsymbol{f} \in R^n \Leftrightarrow \min J_c(\boldsymbol{f}),\ \boldsymbol{f} \in \mathrm{Col}(\boldsymbol{H}). \tag{10}$$

The energy constraint was first obtained for the noiseless case by Johnson and Anderson in [10]. The proofs of Lemma 1 and all subsequent lemmas, and theorems in this paper are all given in the Appendix.

Because of the signal space property, there is a 1 : 1 mapping between the receiver vector $\boldsymbol{f}$ in $\mathrm{Col}(\boldsymbol{H})$ and the combined channel-receiver $\boldsymbol{q} \stackrel{\Delta}{=} \boldsymbol{H}^t\boldsymbol{f}$ in $\mathrm{Row}(\boldsymbol{H})$, as shown in Fig. 2. Therefore, the minimization of $J_c(\boldsymbol{f})$ in $\mathrm{Col}(\boldsymbol{H})$ is equivalent to the minimization of

$$J(\boldsymbol{q}) \stackrel{\Delta}{=} J_c(\boldsymbol{f})|_{\boldsymbol{f}=(\boldsymbol{H}^t)^\dagger \boldsymbol{q}} \tag{11}$$

$$= 3\|\boldsymbol{q}\|_{\boldsymbol{\Phi}}^4 - 2r\|\boldsymbol{q}\|_{\boldsymbol{\Phi}}^2 - (3-r)\|\boldsymbol{q}\|_4^4 + r^2 \tag{12}$$

where, using the fact that $\boldsymbol{q} \in \mathrm{Row}(\boldsymbol{H})$ and the property of pseudo-inverse $\boldsymbol{H}^t(\boldsymbol{H}^t)^\dagger \boldsymbol{q} = \boldsymbol{q}$, we have

$$\boldsymbol{\Phi} = \boldsymbol{I}_m + \sigma^2 \boldsymbol{H}^\dagger (\boldsymbol{H}^\dagger)^t. \tag{13}$$

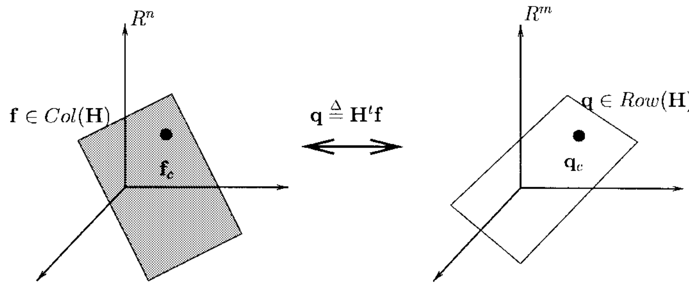


Fig. 2. Equivalent cost function.

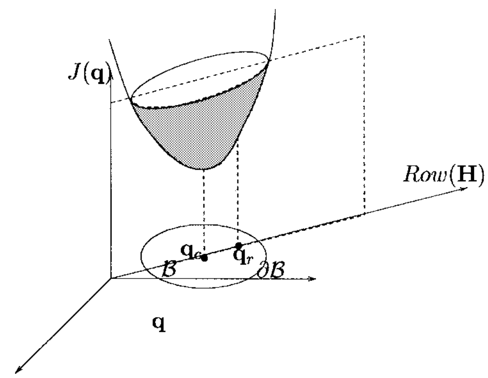


Fig. 3. Geometrical approach with subspace constraint.

According to Lemma 1, CM receivers can be analyzed using the equivalent cost function $J(\boldsymbol{q})$ in $\mathrm{Row}(\boldsymbol{H})$, i.e.,

$$\min J_c(\boldsymbol{f}),\ \boldsymbol{f} \in R^n \ \Leftrightarrow\ \min J(\boldsymbol{q}),\ \boldsymbol{q} \in \mathrm{Row}(\boldsymbol{H}). \tag{14}$$

In contrast to the analysis given in [21], where it is assumed that $\boldsymbol{H}$ has full column rank [hence $\mathrm{Row}(\boldsymbol{H}) = R^m$], the constrained optimization is more general and somewhat more challenging.

1) Geometrical Approach to Locating Minima: Since the evaluation of the gradient and Hessian of the CM cost function is complicated, a geometrical approach is used in this paper to locate CM local minima. The basic idea is to obtain a region, as small as possible, that contains CM receivers defined as local minima of the CM cost function. Suppose that CM receivers are constrained in the linear subspace $\mathrm{Row}(\boldsymbol{H})$ shown in Fig. 3. Suppose that there is a bounded open set $\mathcal{B}$ with boundary $\partial\mathcal{B}$, and $\boldsymbol{q}_r$ is an interior reference point in $\mathcal{B} \bigcap \mathrm{Row}(\boldsymbol{H})$. If the cost $J(\boldsymbol{q})$ on $\partial\mathcal{B} \bigcap \mathrm{Row}(\boldsymbol{H})$ is greater than that of the reference $\boldsymbol{q}_r$, then there exists at least one CM receiver in $\mathcal{B} \bigcap \mathrm{Row}(\boldsymbol{H})$. The principle of this approach is based on the following two points: i) According to the Weierstrass theorem [16, p. 40], there exists a minimum in the compact set $(\mathcal{B} \bigcup \partial\mathcal{B}) \bigcap \mathrm{Row}(\boldsymbol{H})$, and ii) if the CM costs on the boundary are greater than that of the interior reference, there is a minimum inside the region $\mathcal{B} \bigcap \mathrm{Row}(\boldsymbol{H})$.

When the channel is nonsingular ($\mathrm{Row}(\boldsymbol{H}) = \mathcal{R}^m$), this approach is identical to that in [19] and [21]. When the channel is singular ($\mathrm{Row}(\boldsymbol{H}) \neq \mathcal{R}^m$), the difficulty is the constrained optimization of (14). The analyzes based on the nonsingularity of the channel matrix [14], [19], [21] cannot be applied directly. Note that a similar idea of geometric proof has been used by Li *et al.* [15] in a special case. For an

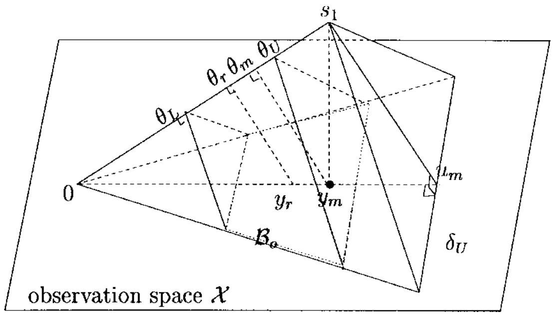


Fig. 4. Neighborhood $\mathcal{B}_o$ in the Hilbert space of the observations.

autoregressive channel model, Li shows that there exist CMA local minima for a finite-length T-spaced equalizer [15]. In comparison with existing results, the main difference is that the approach presented in this paper applies to arbitrary channel models with additive Gaussian noise.

B. Location and MSE Bound of CM Receivers

Our main theorems about the location of CM receivers and their MSE are derived following the three steps in the geometrical approach:

1) Select a neighborhood $\mathcal{B}$.
2) Select a reference $\boldsymbol{q}_r$.
3) Compare $J(\boldsymbol{q})$ on $\partial\mathcal{B}$ with $J(\boldsymbol{q}_r)$.

These steps are described separately below.

1) The Neighborhood: The neighborhood is defined according to the receiver gain and its extra unbiased mean square error (UMSE). For a receiver $\boldsymbol{f}$ that estimates s_1, the receiver gain θ and the (conditionally) unbiased MSE (UMSE) are given by

$$\theta \triangleq \boldsymbol{f}^t\boldsymbol{H}\boldsymbol{e}_1,\ \text{UMSE} = E\left(\frac{1}{\theta}\boldsymbol{f}^t\boldsymbol{x} - s_1\right)^2. \tag{15}$$

Note that $(1/\theta)\boldsymbol{f}^t\boldsymbol{x}$ is a conditionally unbiased estimate of s_1 in the sense that

$$E\left(\frac{1}{\theta}\boldsymbol{f}^t\boldsymbol{x}|s_1\right) = s_1. \tag{16}$$

The geometry involving the linear estimation of s_1 based on $\boldsymbol{x}$ is shown in Fig. 4. The output of any linear estimator must be on the plane $\mathcal{X}$ spanned by the components of $\boldsymbol{x}$. The output y_m of the Wiener receiver is obtained by projecting s_1 on $\mathcal{X}$. If we scale y_m to u_m such that the projection of u_m in the direction of s_1 is s_1, we obtain the so-called (conditionally) unbiased minimum mean square error (U-MMSE) estimate of s_1. Indeed, u_m is conditionally unbiased, i.e., $E(u_m|s_\nu) = s_\nu$. Further, it is recognizable from Fig. 4 that u_m has the shortest distance (and hence the minimum MSE) among all conditionally unbiased estimates. Note that the output of a conditionally unbiased estimator must be on line $\overline{AB}$ due to the orthogonality among sources and noise.

A neighborhood of estimates whose receiver gains (obtained by projecting the estimate in the direction of s_ν) are bounded in (θ_L, θ_U) is shown in the shaded area in Fig. 4, and their corresponding conditionally unbiased estimates of s_1 have mean square error no greater than δ_U^2 over that of u_m. In other words, these estimates have extra (conditionally) unbiased MSE (UMSE) upper bounded by δ_U^2. In this figure, y_r is the output of the reference receiver described later in Section II-B2.

To define this neighborhood mathematically, let the combined channel-receiver $\boldsymbol{q} = \boldsymbol{H}^t\boldsymbol{f}$ have the following parameterization:

$$\begin{aligned}\boldsymbol{q}^t &= [q_1 \cdots q_{\nu-1} q_\nu q_{\nu+1} \cdots q_m]\\ \theta &\triangleq q_\nu = \boldsymbol{e}_\nu^t\boldsymbol{q},\ \boldsymbol{q}_I \triangleq [q_1 \cdots q_{\nu-1} q_{\nu+1} \cdots q_m]/q_\nu.\end{aligned} \tag{17}$$

The receiver output y can be expressed by

$$y = \underbrace{\theta}_{\text{gain}} s_\nu + \underbrace{\sum_{i\neq\nu} q_i s_i}_{\text{interference}} + \underbrace{\boldsymbol{f}^t\boldsymbol{w}}_{\text{noise}} \tag{18}$$

where θ is the receiver gain. Scaling y by $1/\theta$, we have the (conditionally) unbiased estimate of s_ν

$$u \triangleq \frac{y}{\theta} = s_\nu + \underbrace{\frac{1}{\theta}\left(\sum_{i\neq\nu} q_i s_i + \boldsymbol{f}^t\boldsymbol{w}\right)}_{w'\text{: equivalentnoise}}. \tag{19}$$

Therefore, the receiver gain and UMSE of $\boldsymbol{q}$ is given by θ and MSE(u), respectively. Hence, the shaded neighborhood in Fig. 4 is defined by

$$\mathcal{B}_o \triangleq \{\boldsymbol{q}\text{: } \theta_L < \theta < \theta_U,\ \text{MSE}(u) - \text{MSE}(u_m) < \delta_U^2\}. \tag{20}$$

In this definition, θ_L (θ_U) specifies the lower (upper) bound of the CM receiver gain, and δ_U^2 is the upper bound of extra UMSE (see Fig. 4).

Although the neighborhood defined above is specified by particular characteristics of a receiver (UMSE and bias), its relation with the receiver coefficient vector, or equivalently $\boldsymbol{q}$, is not given explicitly. To locate the CM receiver using this neighborhood, it is necessary to translate the above neighborhood to one that is specified by the channel/equalizer parameter space. For this purpose, we introduce the following lemma.

Lemma 2: Let θ, $\boldsymbol{q}_I$, ans u be the gain, interference, and the unbiased receiver output of the receiver $\boldsymbol{q}$. Similar notation with subscript m is defined for the MMSE receiver $\boldsymbol{q}_m$. Let C be the submatrix of $\boldsymbol{\Phi}$ defined in (13) by deleting the νth column and row

$$\begin{aligned}\boldsymbol{\Phi} &= \begin{pmatrix}\boldsymbol{C}_{11} & \boldsymbol{b}_1 & \boldsymbol{C}_{12}\\ \boldsymbol{b}_1^t & a & \boldsymbol{b}_2^t\\ \boldsymbol{C}_{12} & \boldsymbol{b}_2 & \boldsymbol{C}_{22}\end{pmatrix}\\ \boldsymbol{b} &\triangleq \begin{pmatrix}\boldsymbol{b}_1\\ \boldsymbol{b}_2\end{pmatrix}\\ \boldsymbol{C} &\triangleq \begin{pmatrix}\boldsymbol{C}_{11} & \boldsymbol{C}_{12}\\ \boldsymbol{C}_{12} & \boldsymbol{C}_{22}\end{pmatrix}.\end{aligned} \tag{21}$$

Then, in Row$(\boldsymbol{H})$

$$\begin{aligned}&\{\boldsymbol{q}\text{: } \theta_L < \theta < \theta_U,\ \text{MSE}(u) - \text{MSE}(u_m) < \delta_U^2\}\\ &\quad = \{\boldsymbol{q}\text{: } \theta_L < \theta < \theta_U,\ \|\boldsymbol{q}_I - \boldsymbol{q}_{mI}\|_{\boldsymbol{C}} < \delta_U\}\\ &\quad \triangleq \mathcal{B}(\boldsymbol{q}_m, \delta_U, \theta_L, \theta_U).\end{aligned} \tag{22}$$

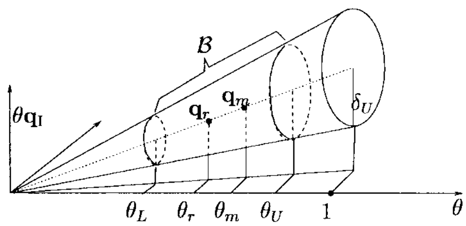


Fig. 5. Cone-type region $\mathcal{B}$.

The equivalence of the two neighborhoods enables us to locate the CM receiver coefficient $\boldsymbol{q}_c$ in the combined channel-receiver space. In Fig. 5, we show that $\mathcal{B}(\boldsymbol{q}_m, \delta_U, \theta_L, \theta_U)$ is a slice of a cone specified by the extra UMSE and receiver gain.

2) The Reference: As shown in Fig. 5, the reference $\boldsymbol{q}_r$ is defined by the vector that is colinear with the MMSE receiver and has the minimum CM cost. Specifically, in estimating s_ν, the MSE of receiver output $y = \boldsymbol{f}^t\boldsymbol{x}$ is given by

$$\begin{aligned}\text{MSE}(y) &= E\{(y - s_\nu)^2\} = E\{(\boldsymbol{f}^t\boldsymbol{x} - s_\nu)^2\} \\ &= \boldsymbol{f}^t\boldsymbol{R}\boldsymbol{f} - 2\boldsymbol{e}_\nu^t\boldsymbol{H}^t\boldsymbol{f} + 1. \end{aligned} \tag{23}$$

The MMSE receiver is then given by

$$\boldsymbol{f}_m = \arg\min_{\boldsymbol{f}} \text{MSE}(\boldsymbol{f}^t\boldsymbol{x}) = \boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{e}_\nu \tag{24}$$

$$\boldsymbol{q}_m = \boldsymbol{H}^t\boldsymbol{f}_m = \boldsymbol{H}^t\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{e}_\nu. \tag{25}$$

Define the reference $\boldsymbol{q}_r = \alpha_r\boldsymbol{q}_m$, where α_r minimizes the CM cost function (12)

$$\begin{aligned}\alpha_r &= \arg\min_\alpha J(\alpha\boldsymbol{q}_m) \\ &= \sqrt{\frac{r\|\boldsymbol{q}_m\|_{\boldsymbol{\Phi}}^2}{3\|\boldsymbol{q}_m\|_{\boldsymbol{\Phi}}^4 - (3-r)\|\boldsymbol{q}_m\|_4^4}}. \end{aligned} \tag{26}$$

The reference $\boldsymbol{q}_r$ should be inside $\mathcal{B}$. This imposes the condition that $\theta_L < \theta_r < \theta_U$.

3) Location of CM Receivers: As mentioned earlier, the key of our approach is to find the neighborhood such that the CM cost on the boundary is uniformly greater than the CM cost at the reference. Having defined the neighborhood $\mathcal{B}(\boldsymbol{q}_m, \delta_U, \theta_L, \theta_U)$ and the reference $\boldsymbol{q}_r$, we are now ready to locate CM receivers by selecting the range of the receiver gains (θ_L, θ_U) and the upper bound δ_c of extra UMSE so that we can prove the necessary inequality. We begin by giving the following lemma, which plays a key role in our approach.

Lemma 3: Let $\boldsymbol{b}$ and $\boldsymbol{C}$ be defined in (21). For all $\boldsymbol{q} \in \mathcal{B}$

$$\begin{aligned}J(\boldsymbol{q}) - J(\boldsymbol{q}_r) &\geq c_2(\delta)\theta^4 + c_1(\delta)\theta^2 + c_0 \\ &\quad(\text{equality holds iff } \delta = 0)\end{aligned} \tag{27}$$

where

$$\delta \triangleq \sqrt{\|\boldsymbol{q}_I - \boldsymbol{q}_{mI}\|_{\boldsymbol{C}}^2 + \delta_0^2} \tag{28}$$

$$\delta_0 \triangleq \|\boldsymbol{q}_{mI} - \boldsymbol{q}_{oI}\|_{\boldsymbol{C}} \tag{29}$$

$$c_0 \triangleq r^2 - J(\boldsymbol{q}_r) \tag{30}$$

$$c_1(\delta) \triangleq -2r\left(\delta^2 + \frac{1}{\theta_o}\right) \tag{31}$$

$$\begin{aligned}c_2(\delta) \triangleq{}& 3\left(\delta^2 + \frac{1}{\theta_o}\right)^2 \\ &- (3-r)(1 + (\delta + \|\boldsymbol{q}_{oI}\|_4)^4)\end{aligned} \tag{32}$$

$$\boldsymbol{q}_{oI} = -\boldsymbol{C}^{-1}\boldsymbol{b} \tag{33}$$

$$\theta_o = \frac{1}{a - \boldsymbol{b}^t\boldsymbol{C}^{-1}\boldsymbol{b}}. \tag{34}$$

According to this lemma, the CM cost function $J(\boldsymbol{q})$ can be reduced to a function in terms of gain θ and extra UMSE δ. Thus, the cone-type region clarifies the CM cost evaluation.

From Lemma 3, it can be seen that the $J(\boldsymbol{q}) - J(\boldsymbol{q}_r)$ is lower bounded by a second-order polynomial of θ^2 with coefficients $c_2(\delta)$, $c_1(\delta)$, and c_0, all of which are functions of δ but not of θ. The region $\mathcal{B}(\boldsymbol{q}_m, \delta_U, \theta_L, \theta_U)$ is obtained by choosing θ_L, θ_U, and δ_U such that $J(\boldsymbol{q}) - J(\boldsymbol{q}_r) > 0$ for all $\boldsymbol{q} \in \partial\mathcal{B} \bigcap \text{Row}(\boldsymbol{H})$. If such θ_L, θ_U, and δ_U exist, then there exists at least one CMA local minimum.

Theorem 1: Given $\boldsymbol{H}$, r, σ^2 and ν with parameters defined in (29)–(32), let $D(\delta) \triangleq c_1^2(\delta) - 4c_2(\delta)c_0$. If

1) $D(\delta)$ has real roots in (δ_0, ∞), the smallest of which is δ_*;
2) $D(\delta_0) \geq 0$;
3) $\forall\,\delta \in [\delta_0, \delta_*]$, $c_2(\delta) > 0$;

then there exists a CM local minimum in

$$\begin{aligned}&\mathcal{B}(\boldsymbol{q}_m, \delta_U, \theta_L, \theta_U) \\ &\quad= \{\boldsymbol{q}\colon \theta_L < \theta < \theta_U, \text{MSE}(u) - \text{MSE}(u_m) < \delta_U^2\}\end{aligned}$$

where

$$\delta_U^2 = \delta_*^2 - \delta_0^2$$

$$\theta_L = \min_{\delta_0 \leq \delta \leq \delta_*} \sqrt{\frac{-c_1(\delta) - \sqrt{c_1^2(\delta) - 4c_2(\delta)c_0}}{2c_2(\delta)}}$$

$$\theta_U = \max_{\delta_0 \leq \delta \leq \delta_*} \sqrt{\frac{-c_1(\delta) + \sqrt{c_1^2(\delta) - 4c_2(\delta)c_0}}{2c_2(\delta)}}.$$

Given the channel matrix $\boldsymbol{H}$, the above theorem enables us i) to test the existence of CM local minima and ii) to obtain the neighborhood containing CM local minima. Further, it provides the bound of extra UMSE and the range of the CM receiver gain.

4) The MSE of CM Receivers: Once $\delta_U, \theta_L, \theta_U$ are obtained from Theorem 1, we can derive the MSE upper bound of CM receivers in this region. We shall see further that because the size of the neighborhood is minimized, the reference $\boldsymbol{q}_r$ turns out to be an accurate approximation of the local minimum in the neighborhood. Therefore, the MSE of the reference is a good estimate of the MSE of the CM receiver. We summarize the MSE bounds and the approximate MSE for the CM receiver in $\mathcal{B}(\boldsymbol{q}_m, \delta_U, \theta_L, \theta_U)$.

Given $\mathbf{H}_{n\times m}$, r, σ^2, ν, MSE upper bound of CM receivers is computed as following:

<u>*Covariance matrix:*</u>

$$\mathbf{R} = \mathbf{H}\mathbf{H}^t + \sigma^2\mathbf{I}_n$$

$$\mathbf{\Phi} = \mathbf{I}_m + \sigma^2(\mathbf{H}^t\mathbf{H})^\dagger$$

$$a = \mathbf{\Phi}(\nu,\nu)$$

$$\mathbf{b} = \mathbf{\Phi}(:,\nu)$$

$\mathbf{C}$ is the part of $\mathbf{\Phi}$ by removing ν-th column and row

<u>*Reference:*</u>

$$\mathbf{q}_m = \mathbf{H}^t\mathbf{R}^{-1}\mathbf{H}\mathbf{e}_\nu$$

$$\alpha_r = \sqrt{\frac{r||\mathbf{q}_m||^2_{\mathbf{\Phi}}}{3||\mathbf{q}_m||^4_{\mathbf{\Phi}} - (3-r)||\mathbf{q}_m||^4_4}}$$

$$\mathbf{q}_r = \alpha_r\mathbf{q}_m$$

$$J(\mathbf{q}_r) = 3||\mathbf{q}_r||^4_{\mathbf{\Phi}} - 2r||\mathbf{q}_r||^2_{\mathbf{\Phi}} - (3-r)||\mathbf{q}_r||^4_4 + r^2$$

$$\mathbf{q}_{o\mathrm{I}} = -\mathbf{C}^{-1}\mathbf{b}, \quad \theta_o = \frac{1}{a - \mathbf{b}^t\mathbf{C}^{-1}\mathbf{b}}$$

$$\delta_0 = ||\mathbf{q}_{m\mathrm{I}} - \mathbf{q}_{o\mathrm{I}}||_{\mathbf{C}}$$

<u>*Condition 1:*</u>

$$\{\delta_1 < \cdots < \delta_k\} = \text{real roots } \{c_1^2(\delta) - 4c_2(\delta)c_0\}$$

$$\delta_* = \min_{\delta_i > \delta_0}\{\delta_1 < \cdots < \delta_k\}$$

If $\delta_* = \emptyset$, can't determine local minima, stop.

<u>*Condition 2:*</u>

If $D(\delta_0) \geq 0$ and $c_2(\delta) > 0 \ \forall\delta \in [\delta_0, \delta_*]$, local minimum exists.

Otherwise, can't determine local minima.

<u>*Gains:*</u>

$$\theta_L = \min_{\delta_0\leq\delta\leq\delta_*}\sqrt{\frac{-c_1(\delta) - \sqrt{c_1^2(\delta) - 4c_2(\delta)c_0}}{2c_2(\delta)}}$$

$$\theta_U = \max_{\delta_0\leq\delta\leq\delta_*}\sqrt{\frac{-c_1(\delta) + \sqrt{c_1^2(\delta) - 4c_2(\delta)c_0}}{2c_2(\delta)}}$$

<u>*MSE:*</u>

$$\mathcal{E}_U = \max\{\frac{(\theta_U - \theta_o)^2}{\theta_o} + (\theta_U\delta_*)^2, \frac{(\theta_L - \theta_o)^2}{\theta_o} + (\theta_L\delta_*)^2\} + 1 - \theta_o.$$

Fig. 6. Algorithm to compute MSE upper bound of CM receivers.

Theorem 2: The MSE of CM receivers in $\mathcal{B}$ is upper bounded by $\mathcal{E}_U$ and is approximated by $\hat{\mathcal{E}}$

$$\mathcal{E}_U = \max\left\{\frac{(\theta_U - \theta_o)^2}{\theta_o} + (\theta_U\delta_*)^2, \frac{(\theta_L - \theta_o)^2}{\theta_o} + (\theta_L\delta_*)^2\right\} + 1 - \theta_o \quad (35)$$

$$\hat{\mathcal{E}} = \frac{(\theta_r - \theta_o)^2}{\theta_o} + (\theta_r\delta_0)^2 + 1 - \theta_o. \quad (36)$$

To assess the quality of the MSE bound, we consider a special case when $\boldsymbol{H}$ has full column rank, and $r = 1$. We are particularly interested in relating the MSE and the extra UMSE bounds to the interference and MSE of the Wiener receiver.

Property 1: Suppose that $\boldsymbol{H}$ is full column rank and that $r = 1$. Let $I_m \triangleq ||\boldsymbol{q}_{mI}||_4$ be the parameter that measures the residual interference. Then

$$\delta_U \approx 2I_m^3 \quad (37)$$

$$\mathcal{E}_U = \mathcal{E}_c + O(\mathcal{E}_m^{2.5}) \quad (38)$$

$$\mathcal{E}_c = \mathcal{E}_m + \tfrac{9}{4}\mathcal{E}_m^2 + O(\mathcal{E}_m^3) \quad (39)$$

$$\boldsymbol{q}_c = \alpha\boldsymbol{q}_m + O(I_m^3) \quad \alpha \text{ is a scale factor.} \quad (40)$$

From (37), because δ_U is the radius of the cone that specified the CM neighborhood, we conclude that for those Wiener receivers with small interference, the CM equalizer is roughly colinear to the MMSE equalizer. This is further demonstrated in (40). The colinear property provides support for using the reference $\boldsymbol{q}_r$ to approximate the true CM receiver because $\boldsymbol{q}_r$ is obtained by minimizing the CM cost in the direction of Wiener receiver. Furthermore, this also implies that the CM receiver will have similar BER performance as that of the MMSE equalizer. Equation (38) shows that the upper bound obtained in Theorem 2 is rather tight, especially for those CM receivers whose corresponding Wiener receiver has small MSE.

Finally, we summarize in Fig. 6 an algorithm that can be used to test the existence of CM receivers and evaluate their locations and MSE performances.

IV. CONCLUSION

In this paper, a MSE upper bound on constant modulus (CM) receiver performance has been derived for an arbitrary channel matrix and Gaussian channel noise. A sufficient condition was given for the existence of a CM receiver in the neighborhood of a Wiener receiver. If such a CM receiver exists and the channel matrix is nonsingular, the extra MSE of the CM receiver has been shown to be the order of the MMSE squared, which implies that the blind receiver design based on the CM criterion achieves almost the same performance as the optimal linear receiver designed for modest amounts of noise. In addition, it has been shown that the unbiased CM receiver vector is almost colinear to the unbiased MMSE receiver vector, which implies that the minimum probability of detection error for linear receivers can be nearly achieved by the CM criterion.

The analysis in this paper is for the static behavior of the CM criterion, which describes the asymptotic achievable performance. An interesting complementary effort would be the study of the dynamic behavior of a CM receiver, focusing, e.g., on the convergence rate and efficient initialization methods.

Appendix

Proof of Lemma 1: First, we prove the energy constraint on CM receiver output. For any $\boldsymbol{f}$ such that $||\boldsymbol{f}||_{\boldsymbol{R}} = 1$, define

$$\phi(k, \boldsymbol{f}) \stackrel{\Delta}{=} J_c\Big(\sqrt{k}\boldsymbol{f}\Big) = \big(3-(3-r)||\boldsymbol{H}^t\boldsymbol{f}||_4^4\big)k^2 - 2rk + r^2. \quad (41)$$

The minimum of $\phi(k, \boldsymbol{f})$ is achieved at

$$k_{\min} = \frac{r}{3-(3-r)||\boldsymbol{H}^t\boldsymbol{f}||_4^4}. \quad (42)$$

Since $0 \le ||\boldsymbol{H}^t\boldsymbol{f}||_4^4 \le ||\boldsymbol{H}^t\boldsymbol{f}||_2^4 \le ||\boldsymbol{f}||_{\boldsymbol{R}}^4 = 1$, $r/3 \le k_{\min} \le 1$. Therefore, $r/3 \le ||\boldsymbol{f}_c||_{\boldsymbol{R}}^2 \le 1$. Furthermore, if $\boldsymbol{H}^t\boldsymbol{f}_c \neq 0$, then $r/3 < ||\boldsymbol{f}_c||_{\boldsymbol{R}}^2 \le 1$.

Second, we derive the CM cost function from a subspace representation. Let $\boldsymbol{U}$ and $\boldsymbol{V}$ be the orthonormal bases of the column space $\boldsymbol{H}$ and its complementary subspace, respectively. Thus, for all $\boldsymbol{f} \in \mathcal{R}^n$, we have $\boldsymbol{f} = \boldsymbol{U}\boldsymbol{u} + \boldsymbol{V}\boldsymbol{v}$. The cost function can be written as

$$\begin{aligned} J_c(\boldsymbol{f}) &= 3\Big(||\boldsymbol{f}||_{\boldsymbol{R}}^2 - \frac{r}{3}\Big)^2 - (3-r)||\boldsymbol{H}^t\boldsymbol{f}||_4^4 + \frac{2r^2}{3} && (43)\\ &= 3\Big(||\boldsymbol{f}||_{\boldsymbol{R}}^2 - \frac{r}{3}\Big)^2 - (3-r)||\boldsymbol{H}^t\boldsymbol{U}\boldsymbol{u}||_4^4 + \frac{2r^2}{3} && (44)\\ &= 3\Big(\Big(||\boldsymbol{U}\boldsymbol{u}||_{\boldsymbol{R}}^2 - \frac{r}{3}\Big) + ||\boldsymbol{V}\boldsymbol{v}||_{\boldsymbol{R}}^2\Big)^2 \\ &\quad - (3-r)||\boldsymbol{H}^t\boldsymbol{U}\boldsymbol{u}||_4^4 + \frac{2r^2}{3} && (45)\\ &= J_c(\boldsymbol{U}\boldsymbol{u}) + 3||\boldsymbol{V}\boldsymbol{v}||_{\boldsymbol{R}}^4 \\ &\quad + 6||\boldsymbol{V}\boldsymbol{v}||_{\boldsymbol{R}}^2\Big(||\boldsymbol{U}\boldsymbol{u}||_{\boldsymbol{R}}^2 - \frac{r}{3}\Big). && (46) \end{aligned}$$

If $\boldsymbol{f} = \boldsymbol{U}\boldsymbol{u} + \boldsymbol{V}\boldsymbol{v}$ is a stationary point of $J_c(\boldsymbol{f})$, $\boldsymbol{f} \in \mathcal{R}^n$, then $\partial J_c(\boldsymbol{f})/\partial\boldsymbol{u} = 0$, and $\partial J_c(\boldsymbol{f})/\partial\boldsymbol{v} = 0$. Note that

$$\begin{aligned} \frac{\partial J_c(\boldsymbol{f})}{\partial \boldsymbol{v}} &= 12||\boldsymbol{V}\boldsymbol{v}||_{\boldsymbol{R}}^2\boldsymbol{v}^H\boldsymbol{V}^H\boldsymbol{V} \\ &\quad + 12\boldsymbol{v}^H\boldsymbol{V}^H\boldsymbol{V}\Big(||\boldsymbol{U}\boldsymbol{u}||_{\boldsymbol{R}}^2 - \frac{r}{3}\Big) && (47)\\ &= 12\boldsymbol{v}^H\Big(||\boldsymbol{V}\boldsymbol{v}||_{\boldsymbol{R}}^2 + ||\boldsymbol{U}\boldsymbol{u}||_{\boldsymbol{R}}^2 - \frac{r}{3}\Big) \\ &= 12\boldsymbol{v}^H\Big(||\boldsymbol{f}||_{\boldsymbol{R}}^2 - \frac{r}{3}\Big). && (48) \end{aligned}$$

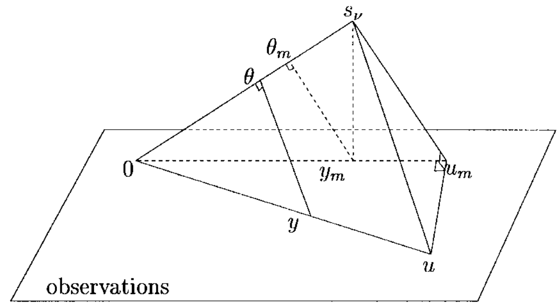


Fig. 7. Hilbert space of the observations.

Thus, $\partial J_c(\boldsymbol{f})/\partial\boldsymbol{v} = 0$ iffy $\boldsymbol{v} = 0$ or $||\boldsymbol{f}||_{\boldsymbol{R}}^2 = r/3$. Therefore,. the stationary point satisfies conditions $\partial J_c(\boldsymbol{f})/\partial\boldsymbol{u} = 0$, $\boldsymbol{v} = 0$, or $\partial J_c(\boldsymbol{f})/\partial\boldsymbol{u} = 0$, $||\boldsymbol{f}||_{\boldsymbol{R}}^2 = r/3$. For the latter case,

$$\frac{\partial J_c(\boldsymbol{f})}{\partial \boldsymbol{u}} = \frac{\partial}{\partial \boldsymbol{u}}\big[-(3-r)||\boldsymbol{H}^t\boldsymbol{U}\boldsymbol{u}||_4^4\big] = 0 \quad (49)$$

implies that $\boldsymbol{u} = 0$. It can be seen that the stationary point satisfying $\boldsymbol{u} = 0$ and $||\boldsymbol{f}||_{\boldsymbol{R}}^2 = r/3$ is a saddle point.

Finally, we show the equivalence of the local minima. Since all CM local minima satisfy the condition $\boldsymbol{v} = 0$, they are in the signal subspace. Therefore, all CM local minima [the minima of $J_c(\boldsymbol{f})$, $\boldsymbol{f} \in R^n$] are local minima of $J_c(\boldsymbol{U}\boldsymbol{u})$ [minima of $J_c(\boldsymbol{f})$, $\boldsymbol{f} \in \text{Col}(\boldsymbol{H})$]. Conversely, if $\boldsymbol{u}_o$ is a minimum of $J_c(\boldsymbol{U}\boldsymbol{u})$, by the result of energy constraint, $||\boldsymbol{U}\boldsymbol{u}_o||_{\boldsymbol{R}}^2 > r/3$. By the definition of local minimum, $\exists\delta > 0$, $\forall||\Delta\boldsymbol{u}||_2 < \delta$, $||\boldsymbol{U}(\boldsymbol{u}_o + \Delta\boldsymbol{u})||_{\boldsymbol{R}}^2 > r/3$, and $J_c(\boldsymbol{U}(\boldsymbol{u}_o + \Delta\boldsymbol{u})) > J_c(\boldsymbol{U}\boldsymbol{u}_o)$. Let $\boldsymbol{f}_o = \boldsymbol{U}\boldsymbol{u}_o$, and $\forall\Delta\boldsymbol{f} = \boldsymbol{U}\Delta\boldsymbol{u} + \boldsymbol{V}\Delta\boldsymbol{v}$ such that $||\Delta\boldsymbol{f}||_2 < \delta$,

$$\begin{aligned} J_c(\boldsymbol{f}_o + \Delta\boldsymbol{f}) &= J_c(\boldsymbol{U}(\boldsymbol{u}_o + \Delta\boldsymbol{u})) + 3||\boldsymbol{V}\Delta\boldsymbol{v}||_{\boldsymbol{R}}^4 \\ &\quad + 6||\boldsymbol{V}\Delta\boldsymbol{v}||_{\boldsymbol{R}}^2\underbrace{\Big(||\boldsymbol{U}(\boldsymbol{u}_o + \Delta\boldsymbol{u})||_{\boldsymbol{R}}^2 - \frac{r}{3}\Big)}_{>0} && (50)\\ &> J_c(\boldsymbol{U}\boldsymbol{u}_o) = J_c(\boldsymbol{f}_o). && (51) \end{aligned}$$

Hence, $\boldsymbol{f}_o$ is a minimum of $J_c(\boldsymbol{f})$, $\boldsymbol{f} \in R^n$. □□□

Proof of Lemma 2: First, the relationship between u and u_m is depicted in Fig. 7. It will be shown that $u_m - u$ is orthogonal to u_m and $s_\nu - u_m$. Since $s_\nu - y_m$ is orthogonal to the subspace of observations, then

$$s_\nu - y_m \perp u - u_m. \quad (52)$$

From the definition of the unbiased estimator, $(u - s_\nu) \perp s_\nu$, and then

$$s_\nu \perp u - u_m. \quad (53)$$

From (52) and (53), $u - u_m \perp y_m$. Since u_m is the scaled y_m, therefore

$$u - u_m \perp u_m \quad (54)$$

$$u - u_m \perp s_\nu - u_m. \quad (55)$$

Based on the above orthogonal properties, $E\{(u - s_\nu)^2\} = E\{(u_m - s_\nu)^2\} + E\{(u - u_m)^2\}$, i.e., $\text{MSE}(u) - \text{MSE}(u_m) =$

$E\{(u-u_m)^2\}$. For $\boldsymbol{q} \in \text{Row}(\boldsymbol{H})$

$$\begin{aligned}\text{MSE}(u) - \text{MSE}(u_m) &= E\{(u-u_m)^2\} \\ &= \left\| \frac{\boldsymbol{q}}{\boldsymbol{e}_\nu^t \boldsymbol{q}} - \frac{\boldsymbol{q}_m}{\boldsymbol{e}_\nu^t \boldsymbol{q}_m} \right\|_{\boldsymbol{\Phi}}^2 \\ &= \|\boldsymbol{q}_I - \boldsymbol{q}_{mI}\|_{\boldsymbol{C}}^2. \end{aligned} \tag{56}$$

□□□

Proof of Lemma 3: For CM cost function (12), there are two terms $\|\boldsymbol{q}\|_{\boldsymbol{\Phi}}^2$ and $\|\boldsymbol{q}\|_4$ that need to be derived. From Fig. 7 and the orthogonal properties (54) and (55)

$$\boldsymbol{q}^t \boldsymbol{\Phi} \boldsymbol{q} = E\{y^2\} = \theta^2 E\{u^2\} = \theta^2(E\{(u-u_m)^2\} + E\{u_m^2\}). \tag{57}$$

Note that

$$\begin{aligned} E\{u_m^2\} &= \left\| \frac{\boldsymbol{q}_m}{\boldsymbol{e}_\nu^t \boldsymbol{q}_m} \right\|_{\boldsymbol{\Phi}}^2 \\ &= (a^2 + 2\boldsymbol{b}^t \boldsymbol{q}_{mI} + \boldsymbol{q}_{mI}^t \boldsymbol{C} \boldsymbol{q}_{mI}) \qquad (58) \\ &= (\boldsymbol{q}_{mI} + \boldsymbol{C}^{-1}\boldsymbol{b})^t \boldsymbol{C} (\boldsymbol{q}_{mI} + \boldsymbol{C}^{-1}\boldsymbol{b}) \\ &\quad + (a - \boldsymbol{b}^t \boldsymbol{C}^{-1} \boldsymbol{b}). \qquad (59) \end{aligned}$$

Since $\boldsymbol{q}_{oI} = -\boldsymbol{C}^{-1}\boldsymbol{b}$ and $\theta_o = 1/(a - \boldsymbol{b}^t\boldsymbol{C}^{-1}\boldsymbol{b})$, then $E\{u_m^2\} = \|\boldsymbol{q}_{mI} - \boldsymbol{q}_{oI}\|_{\boldsymbol{C}}^2 + (1/\theta_o)$, and

$$\begin{aligned} \boldsymbol{q}^t \boldsymbol{\Phi} \boldsymbol{q} &= \theta^2 \left(\|\boldsymbol{q}_I - \boldsymbol{q}_{mI}\|_{\boldsymbol{C}}^2 + \underbrace{\|\boldsymbol{q}_{mI} - \boldsymbol{q}_{oI}\|_{\boldsymbol{C}}^2}_{\delta_o^2} \right) + \frac{\theta^2}{\theta_o} \\ &= \theta^2 \underbrace{(\|\boldsymbol{q}_I - \boldsymbol{q}_{mI}\|_{\boldsymbol{C}}^2 + \delta_0^2)}_{\delta^2} + \frac{\theta^2}{\theta_o} = \theta^2\delta^2 + \frac{\theta^2}{\theta_o}. \end{aligned} \tag{60}$$

For the second term $\|\boldsymbol{q}_I\|_4$, we have

$$\begin{aligned} \|\boldsymbol{q}_I\|_4 &= \|\boldsymbol{q}_I - \boldsymbol{q}_{oI} + \boldsymbol{q}_{oI}\|_4 \\ &\le \|\boldsymbol{q}_I - \boldsymbol{q}_{oI}\|_4 + \|\boldsymbol{q}_{oI}\|_4 \\ &\le \|\boldsymbol{q}_I - \boldsymbol{q}_{oI}\|_{\boldsymbol{C}} + \|\boldsymbol{q}_{oI}\|_4 \\ &= \delta + \|\boldsymbol{q}_{oI}\|_4. \end{aligned} \tag{61}$$

Substituting above two terms into (12), we have

$$\begin{aligned} &J(\boldsymbol{q}) - J(\boldsymbol{q}_r) \\ &= 3(\boldsymbol{q}^t\boldsymbol{\Phi}\boldsymbol{q})^2 - 2r(\boldsymbol{q}^t\boldsymbol{\Phi}\boldsymbol{q}) - (3-r)\|\boldsymbol{q}\|_4^4 + r^2 - J(\boldsymbol{q}_r) \\ &\ge 3\left(\theta^2\delta^2 + \frac{\theta^2}{\theta_o}\right)^2 \theta^4 - 2r\left(\theta^2\delta^2 + \frac{\theta^2}{\theta_o}\right)\theta^2 \\ &\quad - (3-r)\theta^4(1 + (\delta + \|\boldsymbol{q}_{oI}\|_4)^4) + r^2 - J(\boldsymbol{q}_r) \\ &= \underbrace{\left(3\left(\delta^2 + \frac{1}{\theta_o}\right)^2 - (3-r)(1 + (\delta + \|\boldsymbol{q}_{oI}\|_4)^4)\right)}_{c_2(\delta)} \theta^4 \\ &\quad \underbrace{-2r\left(\delta^2 + \frac{1}{\theta_o}\right)}_{c_1(\delta)} \theta^2 + \underbrace{r^2 - J(\boldsymbol{q}_r)}_{c_0}. \end{aligned} \tag{62}$$

□□□

Proof of Theorem 1: According to (22), $\mathcal{B}$ is given by $\mathcal{B}(\boldsymbol{q}_m, \delta_U, \theta_L, \theta_U) = \{\boldsymbol{q}: \theta_L < \theta < \theta_U, \|\boldsymbol{q}_I - \boldsymbol{q}_{mI}\|_{\boldsymbol{C}} < \delta_U\}$. The boundary $\partial\mathcal{B}$ consists of

$$\begin{aligned} \partial\mathcal{B}_1 &= \{\boldsymbol{q}: \|\boldsymbol{q}_I - \boldsymbol{q}_{mI}\|_{\boldsymbol{C}} = \delta_U\} \\ &= \{\boldsymbol{q}: \|\boldsymbol{q}_I - \boldsymbol{q}_{mI}\|_{\boldsymbol{C}}^2 + \delta_0^2 = \delta_*^2\} \\ \partial\mathcal{B}_2 &= \{\boldsymbol{q}: \theta = \theta_L \text{ or } \theta = \theta_U \|\boldsymbol{q}_I - \boldsymbol{q}_{mI}\|_{\boldsymbol{C}} \le \delta_U\}. \end{aligned}$$

Next, $J(\boldsymbol{q}) - J(\boldsymbol{q}_r) > 0$ will be proven for all each sub-boundary.

1) According to the conditions of the theorem, $D(\delta_*) = 0$ and $c_2(\delta_*) > 0$. Thus, $c_2(\delta_*)\theta^4 + c_1(\delta_*)\theta^2 + c_0 \ge 0$ for all θ. Since $\delta_* > 0$, from Lemma 3, $J(\boldsymbol{q}) > J(\boldsymbol{q}_r)$ for all points in $\mathcal{B}_1$.
2) For all points in $\text{Row}(\boldsymbol{H})$, $\delta \triangleq \sqrt{\|\boldsymbol{q}_I - \boldsymbol{q}_{mI}\|_{\boldsymbol{C}}^2 + \delta_0^2} \ge \delta_0$. Thus, $\mathcal{B}_2 \bigcap \text{Row}(\boldsymbol{H})$ is equivalent to $\{\boldsymbol{q}: \theta = \theta_L \text{ or } \theta = \theta_U, \delta_0 \le \delta \le \delta_*\}$. For all points in this set, $c_2(\delta) > 0$. Thus, if θ^2 is not in the interval
$$\left[\frac{-c_1(\delta) - \sqrt{c_1^2(\delta) - 4c_2(\delta)c_0}}{2c_2(\delta)} \quad \frac{-c_1(\delta) + \sqrt{c_1^2(\delta) - 4c_2(\delta)c_0}}{2c_2(\delta)} \right] \tag{63}$$
then the polynomial $c_2(\delta)\theta^4 + c_1(\delta)\theta^2 + c_0 > 0$. According to the definition of θ_U and θ_L in the theorem, $J(\boldsymbol{q}) > J(\boldsymbol{q}_r)$ for all points in $\mathcal{B}_2 \bigcap \text{Row}(\boldsymbol{H})$.
3) Finally, we verify that the reference $\boldsymbol{q}_r$ is in the region $\mathcal{B}_p(\boldsymbol{q}_m, \delta_U, \theta_L, \theta_U)$, i.e., $\theta_L < \theta_r < \theta_U$. According to Lemma 3, $c_2(\delta_0)\theta_r^4 + c_1(\delta_0)\theta_r^2 + c_0 \le 0$. Therefore
$$\sqrt{\frac{-c_1(\delta_0) - \sqrt{c_1(\delta_0)^2 - 4c_2(\delta_0)c_0}}{2c_2(\delta_0)}} \le \theta_r \le \sqrt{\frac{-c_1(\delta_0) + \sqrt{c_1(\delta_0)^2 - 4c_2(\delta_0)c_0}}{2c_2(\delta_0)}}. \tag{64}$$
Therefore, $\theta_L < \theta_r < \theta_U$, $\boldsymbol{q}_r$ is in the region $\mathcal{B}$.

□□□

Proof of Theorem 2: From (23) and (60), the MSE of receiver $\boldsymbol{q}$ is given by

$$\begin{aligned} J_m(\boldsymbol{q}) &= \boldsymbol{q}^t\boldsymbol{\Phi}\boldsymbol{q} - 2\boldsymbol{e}_\nu^t\boldsymbol{q} + 1 \\ &= \left(\theta^2\delta^2 + \frac{\theta^2}{\theta_o}\right) - 2\theta + 1 \\ &= \theta^2\delta^2 + \frac{(\theta - \theta_0)^2}{\theta_0} + (1 - \theta_0). \end{aligned} \tag{65}$$

From Theorem 1, for all points in $\mathcal{B} \bigcap \text{Row}(\boldsymbol{H})$, $\delta < \delta_*$, and $\theta_L < \theta < \theta_U$. Therefore

$$\begin{aligned} J_m(\boldsymbol{q}) &\le \theta^2\delta_*^2 + \frac{(\theta - \theta_0)^2}{\theta_0} + (1 - \theta_0) \\ &\le \max_\theta \left(\theta^2\delta_*^2 + \frac{(\theta - \theta_0)^2}{\theta_0}\right) + (1 - \theta_0) \\ &= \max\left\{ \frac{(\theta_U - \theta_o)^2}{\theta_o} + (\theta_U\delta_*)^2, \frac{(\theta_L - \theta_o)^2}{\theta_o} + (\theta_L\delta_*)^2 \right\} + 1 - \theta_o. \end{aligned} \tag{66}$$

If we use the reference to approximate a CM receiver in $\mathcal{B}\bigcap \text{Row}(\boldsymbol{H})$, we have

$$J_m(\boldsymbol{q}_r) = \theta_r^2\delta_0^2 + \frac{(\theta_r-\theta_0)^2}{\theta_0} + (1-\theta_0). \tag{67}$$

□□□

Proof of Property 1: At high SNR, the second-order approximation of $D(\delta)$ is given by

$$\begin{aligned}D(\delta) =& 4\left(\delta^2+\frac{1}{\theta_m}\right)^2 \\ &- 4\frac{3\left(\delta^2+\frac{1}{\theta_m}\right)^2 - 2(1+(\delta+I_m)^4)}{3-2\theta_m^2-2\theta_m^2 I_m^4} \\ =& \frac{32I_m^3}{3-2\theta_m^2-2\theta_m^2 I_m^4}\delta \\ &+ 4\left(\frac{2}{\theta_m} - \frac{6/\theta_m - 12I_m^2}{3-2\theta_m^2-2\theta_m^2 I_m^4}\right)\delta^2 + O(\delta^3) \\ =& \frac{4}{(3-2\theta_m^2-2\theta_m^2 I_m^4)\theta_m} \\ &\cdot (8I_m^3\theta_m\delta + (2(3)-2\theta_m^2-2\theta_m^2 I_m^4) \\ &- (6-12I_m^2\theta_m))\delta^2) + O(\delta^3) \\ \approx& \frac{4}{(3-2\theta_m^2-2\theta_m^2 I_m^4)\theta_m}(8I_m^3\theta_m\delta - 4\theta_m^2\delta^2).\end{aligned}$$

Thus, $\delta_U \approx 2I_m^3/\theta_m \approx 2I_m$. Since $||\boldsymbol{q}_{cI} - \boldsymbol{q}_{mI}||_2 < \delta_U$, then

$$\boldsymbol{q}_c = \alpha\boldsymbol{q}_m + O(I_m^3). \tag{68}$$

For the MSE bound, we need to approximate θ_U and θ_L first. Since $c_1(\delta) = c_1(0)+O(\delta^2) \approx -2/\theta_m$, $c_2(\delta) = c_2(0)+O(\delta)$, we have

$$\begin{aligned}\theta_U =& \max_{0<\delta<\delta_U} \sqrt{\frac{-c_1(\delta)+\sqrt{D(\delta)}}{2c_2(\delta)}} \\ \approx& \max_{0<\delta<\delta_U} \sqrt{\frac{-c_1(0)+\sqrt{D(\delta)}}{2c_2(0)}} \qquad (69)\\ =& \max_{0<\delta<\delta_U} \underbrace{\sqrt{\frac{-c_1(0)}{2c_2(0)}}}_{\theta_r}\sqrt{1-\frac{\sqrt{D(\delta)}}{c_1(0)}} \\ \approx& \max_{0<\delta<\delta_U} \theta_r\left[1-\frac{1}{2c_1(0)}\sqrt{D(\delta)}\right] \qquad (70)\\ \approx& \max_{0<\delta<\delta_U} \theta_r[1+\frac{\theta_m}{4}\sqrt{32I_m^3\delta - 16\delta^2}] \\ =& \theta_r(1+\theta_m I_m^3) \approx \theta_r(1+I_m^3). \qquad (71)\end{aligned}$$

Thus, $\theta_U+\theta_L \approx 2\theta_r$, and $\theta_U-\theta_L \approx 2I_m^3$. Therefore, we can justify the accuracy of the MSE bound by

$$\begin{aligned}\mathcal{E}_U - J_m(\boldsymbol{q}_c) =& \max_{\boldsymbol{q}\in\mathcal{B}} J_m(\boldsymbol{q}) - J_m(\boldsymbol{q}_c) \\ \le& \max_{\boldsymbol{q}\in\mathcal{B}} J_m(\boldsymbol{q}) - \min_{\boldsymbol{q}\in\mathcal{B}} J_m(\boldsymbol{q}) \qquad (72)\\ =& \max_{\theta_L<\theta<\theta_U,\,\delta<\delta_U}\left\{\frac{(\theta-\theta_m)^2}{\theta_m}+(\theta\delta)^2\right\} \\ &- \min_{\theta_L<\theta<\theta_U,\,\delta<\delta_U}\left\{\frac{(\theta-\theta_m)^2}{\theta_m}+(\theta\delta)^2\right\} \\ \le& \frac{|(\theta_U-\theta_m)^2-(\theta_L-\theta_m)^2|}{\theta_m} + (\theta_U\delta_U)^2 \\ \approx& \frac{2|\theta_r-\theta_m|}{\theta_m}(\theta_U-\theta_L) + (\theta_U\delta_U)^2 \\ =& O(\mathcal{E}_m)O(I_m^3) + (O(I_m^3)O(I_m^3))^2. \qquad (73)\end{aligned}$$

Since $I_m^2 < \mathcal{E}_m$, then

$$\mathcal{E}_U - J_m(\boldsymbol{q}_c) < O(\mathcal{E}_m^{2.5}). \tag{74}$$

To show (39), we approximate the CMA solution $\boldsymbol{q}_c$ by the reference $\boldsymbol{q}_r$

$$\begin{aligned}\mathcal{E}_c - \mathcal{E}_m \approx& E\{(y_r-y_m)^2\} \\ =& \frac{(\theta_r-\theta_m)^2}{\theta_m} \\ =& \left(\frac{\theta_r}{\sqrt{\theta_m}} - \sqrt{\theta_m}\right)^2. \qquad (75)\end{aligned}$$

According to (26)

$$\begin{aligned}\frac{\theta_r}{\sqrt{\theta_m}} =& (1+2(1-\theta_m^2(1+||\boldsymbol{q}_{mI}||_4^4)))^{-(1/2)} \\ =& 1-(1-\theta_m^2(1+||\boldsymbol{q}_{mI}||_4^4)) + O(1-\theta_m^2(1+||\boldsymbol{q}_{mI}||_4^4)).\end{aligned} \tag{76}$$

Since

$$\begin{aligned}1-\theta_m^2(1+||\boldsymbol{q}_{mI}||_4^4) =& (1+\theta_m)(1-\theta_m) - \theta_m^2||\boldsymbol{q}_{mI}||_4^4 \\ =& 2J_m(\boldsymbol{q}_m) + O(J_m(\boldsymbol{q}_m))\end{aligned} \tag{77}$$

and thus

$$\begin{aligned}&\frac{\theta_r}{\sqrt{\theta_m}} - \sqrt{\theta_m} \\ &= [1-2J_m(\boldsymbol{q}_m)] - \left[1-\tfrac{1}{2}J_m(\boldsymbol{q}_m)\right] + O(J_m(\boldsymbol{q}_m)).\end{aligned} \tag{78}$$

Therefore

$$\mathcal{E}_c - \mathcal{E}_m = \tfrac{9}{4}\mathcal{E}_m^2 + O(\mathcal{E}_m^3). \tag{79}$$

□□□

ACKNOWLEDGMENT

The authors would like to thank D. R. Brown for his detailed comments on this paper.

References

[1] 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.

[2] 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.

[3] 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.

[4] ——, "Local convergence of the Sato blind equalizer and generalizations under practical constraints," *IEEE Trans. Inform. Theory,* vol. 39, pp. 129–144, Jan. 1993.

[5] I. Fijalkow, A. Touzni, and J. R. Treichler, "Fractionally-spaced equalization by CMA: Robustness to channel noise and lack of disparity," *IEEE Trans. Signal Processing,* vol. 45, pp. 56–66, Jan. 1997.

[6] I. Fijalkow, J. R. Treichler, and C. R. Johnson, Jr., "Fractionally spaced blind equalization: Loss of channel disparity," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process.,* Detroit, MI, May 1995, vol. 3, pp. 1988–1991.

[7] G. J. Foschini, "Equalizing without altering or detecting data," *Bell Syst. Tech. J.,* pp. 64, 1885–1911, Oct. 1985.

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

[9] 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. Acoust., Speech, Signal Processing,* vol. 36, pp. 900–910, June 1988.

[10] C. R. Johnson, Jr. and B. D. O. Anderson, "Godard blind equalizer error surface characteristics: White, zero-mean, binary source case," *Int. J. Adaptive Contr. Signal Processing,* pp. 301–324, 1995.

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

[12] P. Lancaster and M. Tismenetsky, *The Theory of Matrices.* New York: Academic, 1984.

[13] Y. Li and Z. Ding, "Convergence analysis of finite length blind adaptive equalizers," *IEEE Trans. Signal Processing,* vol. 43, pp. 2120–2129, Sept. 1995.

[14] ——, "Global convergence of fractionally spaced Godard (CMA) adaptive equalizers," *IEEE Trans. Signal Processing,* vol. 44, pp. 818–826, Apr. 1996.

[15] Y. Li, 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.

[16] D. G. Luenberger, *Optimization by Vector Space Methods.* New York: Wiley, 1969.

[17] A. Touzni, I. Fijalkow, and J. R. Treichler, "Fractionally-spaced CMA under channel noise," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process.,* Atlanta, GA, May 1996, vol. 5, pp. 2674–2677.

[18] J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-31, pp. 459–472, Apr. 1983.

[19] 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. 27th Conf. Inform. Sci., Syst.,* Baltimore, MD, Mar. 1996, pp. 890–894.

[20] H. Zeng, L. Tong, and C. R. Johnson, "Behavior of fractionally-spaced constant modulus algorithm: Mean square error, robustness and local minima," in *Proc. 30th Asilomar Conf. Signals, Syst., Comput.,* Nov. 1996, vol. II, pp. 305–309.

[21] ——, "Relationships between CMA and wiener receivers," *IEEE Trans. Inform. Theory,* vol. IT-44, Apr. 1998.

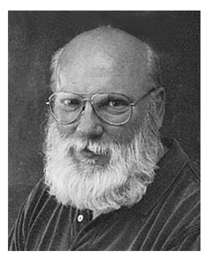
Hanks H. Zeng (M'98) was born in Beijing, China, in 1965. He received the B.E. degree in electrical engineering in applied mathematics from Tsinghua University, Beijing, in 1989, the M.S. degree in acoustics from the Chinese Academy of Science, Nanjing, in 1992, and the Ph.D. degree in electrical engineering from the University of Connecticut, Storrs, in 1997, respectively.

From 1997 to 1999, he worked for Philips Consumer Communications, Piscataway, NJ. Since March 1999, he has been with AT&T Laboratories—Research, Red Bank, NJ. His research interests include equalization techniques, estimation theory, and performance analysis.

Lang Tong (S'87–M'91) received the B.E. degree from Tsinghua University, Beijing, China, in 1985 and the M.S. and Ph.D. degrees in electrical engineering in 1987 and 1990, respectively, from the University of Notre Dame, Notre Dame, IN.

After being a Postdoctoral Research Affiliate at the Information Systems Laboratory, Stanford University, Stanford, CA, he joined the Department of Electrical and Computer Engineering, West Virginia University, Morgantown, and was also with the University of Connecticut, Storrs. Since the fall of 1998, he has been with the School of Electrical Engineering, Cornell University, where he is an Associate Professor. He also held a Visiting Assistant Professor position at Stanford University in the summer of 1992. His research interests include statistical signal processing, wireless communication, and system theory.

Dr. Tong received the Young Investigator Award from the Office of Naval Research in 1996 and the Outstanding Young Author Award from the IEEE Circuits and Systems Society.



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

He is currently a Professor of Electrical 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 recent principal focus for has been blind equalization.