Relationships Between the Constant Modulus and Wiener Receivers

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

***Abstract*— The Godard or the constant modulus algorithm (CMA) is an effective technique for blind receiver design in communications. However, due to the complexity of the Constant Modulus (CM) cost function, the performance of CM receivers has primarily been evaluated using simulations. Theoretical analysis is typically based on either the noiseless case or approximations of the cost function. The following question, while resolvable numerically for a specific example, remains unanswered in a generic manner: *In the presence of channel noise, where are the CM local minima and what are their mean-squared errors (MSE)?* In this paper, a geometrical approach is presented that relates CM to Wiener (or minimum MSE) receivers. Given the MSE and the intersymbol/user interference of a Wiener receiver, a sufficient condition is given for the existence of a CM local minimum in the neighborhood of the Wiener receiver. MSE bounds on CM receiver performance are derived and shown to be tight in simulations. The analysis shows that, while in some cases the CM receiver performs almost as well as the (nonblind) Wiener receiver, it is also possible that, due to its blind nature, CM receiver may perform considerably worse than a (nonblind) Wiener receiver.**

***Index Terms*— Adaptive filters, blind deconvolution, constant modulus algorithm (CMA), equalization, intersymbol interference, local convergence, Wiener receiver.**

I. Introduction

A. The Problem

Linear estimation of a random variable s from observation random vector $\boldsymbol{x}$ is a classical problem. The well-known Wiener receiver, often referred to as the minimum mean-squared error (MMSE) receiver, is obtained by minimizing the mean-squared error (MSE)

$$J_m(\boldsymbol{f}) \triangleq E\{(\boldsymbol{f}^t\boldsymbol{x} - s)^2\}, \quad \boldsymbol{f}_m \triangleq \arg\min_{\boldsymbol{f}} J_m(\boldsymbol{f}). \quad (1)$$

where $\boldsymbol{f}$ is the receiver parameter vector. As an alternative, Godard [8] and Treichler *et al.* [24] proposed the constant modulus (CM) criterion which minimizes the dispersion of the receiver output around the dispersion constant $R_p = E\{|s|^4\}/E\{|s|^2\}$

$$J_c(\boldsymbol{f}) \triangleq E\{(|\boldsymbol{f}^t\boldsymbol{x}|^2 - R_p)^2\}. \quad (2)$$

The locations in receiver parameter space of the local minima of $J_c(\boldsymbol{f})$ are referred to as constant modulus (CM) receivers.

Clearly MSE and CM are different criteria. The Wiener receiver requires the *joint* second-order moment of $\boldsymbol{x}$ and s and has a closed-form solution. When the joint moment is not known in practical applications, training data may be needed. In contrast, the CM receiver is blind because it only requires *marginal* moments of the observation $\boldsymbol{x}$ and source s. However, the CM cost function J_c may have local minima, and CM receivers do not have closed-form solutions.

The main objective of this paper is to reveal connections between CM and Wiener receivers. We aim to answer (partially) the following questions:

- Q1: Does CM have local minima? If so, where are they?
- Q2: What is the mean-squared error (MSE) performance of CM receivers?
- Q3: What is the relationship between (blind) CM and (nonblind) Wiener receivers?

B. An Example

To obtain some insights into the above questions, let us consider an example of $T/2$-spaced equalization used in [20]. The model for the fractionally spaced equalization is equivalent to a multirate system [19], [26] shown in Fig. 1. The up-sampler is an interpolator which inserts a zero between successive symbols $s(k)$, while the down-sampler selects a baud-rate sequence from $\{z(n)\}$. In this model, the transmitted signal $\{s(k)\}$ at symbol rate $1/T$ is a random sequence taking values at ± 1 with equal probability (i.e., uncoded binary phase-shift keying (BPSK) signal). The channel impulse response is $h(0) = 1$, $h(1) = 1$, $h(2) = 3$, $h(3) = 0.5$. The $T/2$-spaced equalizer has two taps, i.e., $\boldsymbol{f} = [f(0), f(1)]^t$. Thus we have

$$x(n) = \sum_k h(n - 2k)s(k) + w(n) \quad (3)$$

$$y(k) \triangleq z(2k+1) = \sum_l f(l)x(2k+1-l)$$
$$= q(0)s(k) + q(1)s(k-1) + n(k) \quad (4)$$

Manuscript received September 6, 1996; revised December 2, 1997. The work of H. Zeng and L. Tong was supported in part by the National Science Foundation under Contract NCR-9321813, by the Office of Naval Research under Contract N00014-96-1-0895, and by the Advanced Research Projects Agency monitored by the Federal Bureau of Investigation under Contract J-FBI-94-221. The work of C. R. Johnson, Jr., was supported in part by the NSF under Grant MIP-9509011 and by Applied Signal Technology. The material in this paper was presented in part at the Conference of Information Sciences and Systems, Princeton, NJ, March 1996.

H. H. Zeng and L. Tong are with the Department of Electrical and Systems Engineering, University of Connecticut, Storrs, CT 06269 USA (e-mail: ltong@eng2.uconn.edu).

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

Publisher Item Identifier S 0018-9448(98)03461-0.

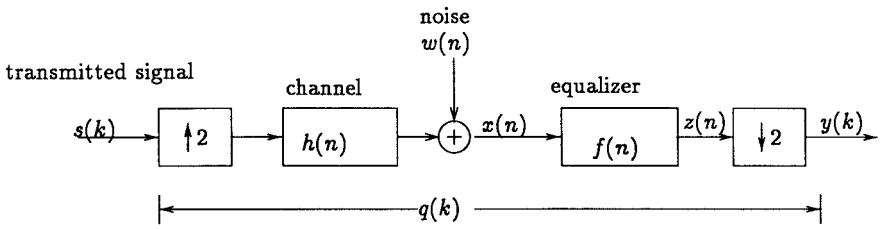

Fig. 1. A $T/2$-spaced equalizer is equivalent to a multirate system.

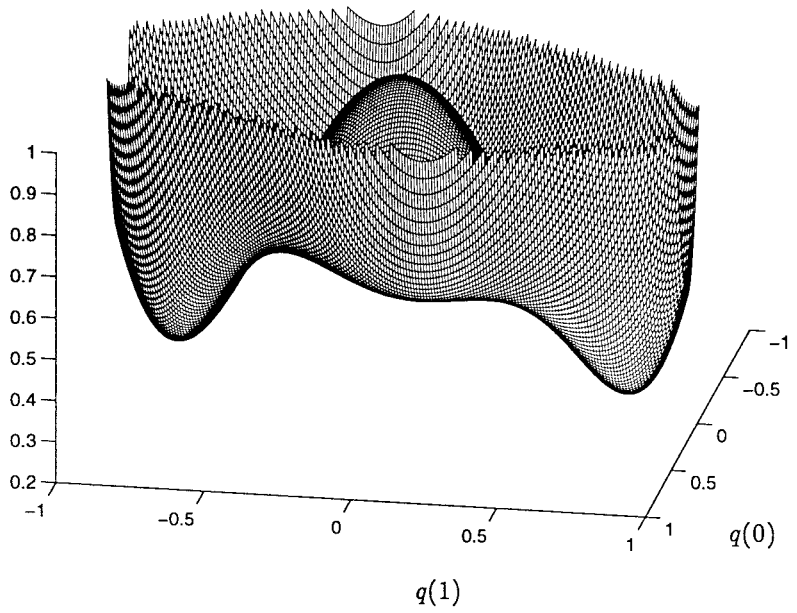

Fig. 2. The CM cost function at SNR = 10 dB.

where $n(k)$ is noise at the down-sampler output, and $q(n)$ is the overall response of the combined channel and receiver. By the definition in [18, p. 603], when $\boldsymbol{q} = [q(0), q(1)]^t = [1, 0]^t$ and $\boldsymbol{q} = [0, 1]^t$, the corresponding $\boldsymbol{f}$'s are zero-forcing (ZF) equalizers for $s(k)$ and $s(k-1)$, respectively.

For BPSK signal, the dispersion constant $R_p = E\{|s|^4\}/E\{|s|^2\} = 1$. According to (2), the CMA receiver is obtained by minimizing

$$J_c(\boldsymbol{f}) = E\{(|y(k)|^2 - 1)^2\} \triangleq J(\boldsymbol{q}) \tag{5}$$

which is shown in Fig. 2. Fig. 3 gives the contours of $J(\boldsymbol{q})$ and the locations of the Wiener and CM receivers. Because $J(\boldsymbol{q})$ is symmetrical, only the two receivers estimating $s(k)$ indicated by $\times$ and $s(k-1)$ indicated by $\otimes$ are considered. These figures suggest the following.

- For Q1 and Q3, we observe that there exists a CM local minimum, and the CM local minima ($\times$ and $\otimes$ in Fig. 3) are close to the Wiener receivers ($+$ and $\oplus$ in Fig. 3).
- For Q2, the MSE's of the CM receivers ($\times$ and $\otimes$ in Fig. 3) are 0.4393 and 0.1225, respectively. Without proper initialization, the CM equalizer may converge to $(\times)$ and have large MSE.

Although the above observations are based on this artificial example, they turn out to be true in general, which will be established analytically in this paper.

C. A Geometrical Approach and Main Results

Our approach to analyzing CM receiver with a finite number of parameters is geometrical, which is in contrast to commonly used methods of analyzing equilibria points of the CM cost function and their local curvatures. The basic idea is based on the Weierstrass maximum theorem [17, p. 40]. As illustrated in Fig. 4, our approach is to find a compact region $\mathcal{B}$ with boundary $\partial\mathcal{B}$ in the parameter space, and an interior reference $\boldsymbol{f}_r$ such that the CM costs $J_c(\boldsymbol{f})$ on the boundary $\partial\mathcal{B}$ are greater than that of the reference $J_c(\boldsymbol{f}_r)$. Consequently, there must be at least a local minimum of CM in $\mathcal{B}$.

Critical in this approach is the selection of the shape of $\mathcal{B}$, the location of $\mathcal{B}$, and the reference $\boldsymbol{f}_r$. In defining such a region, our first goal is to have it as small as possible, which leads to a more accurate description of the local minimum and its MSE performance. The second goal is to relate such a region to the location of Wiener receivers. In addition to the description of a $\mathcal{B}$ in the parameter space, we aim to give a corresponding description in the Hilbert space of the observations in order to provide insightful physical interpretations.

While there are important differences between CMA in an equalization problem for a single user and in beamforming for multiuser (see [9]), the problem of fractionally spaced equalization of intersymbol interference channels and the problem of beamforming, i.e., separating the signal of interests from

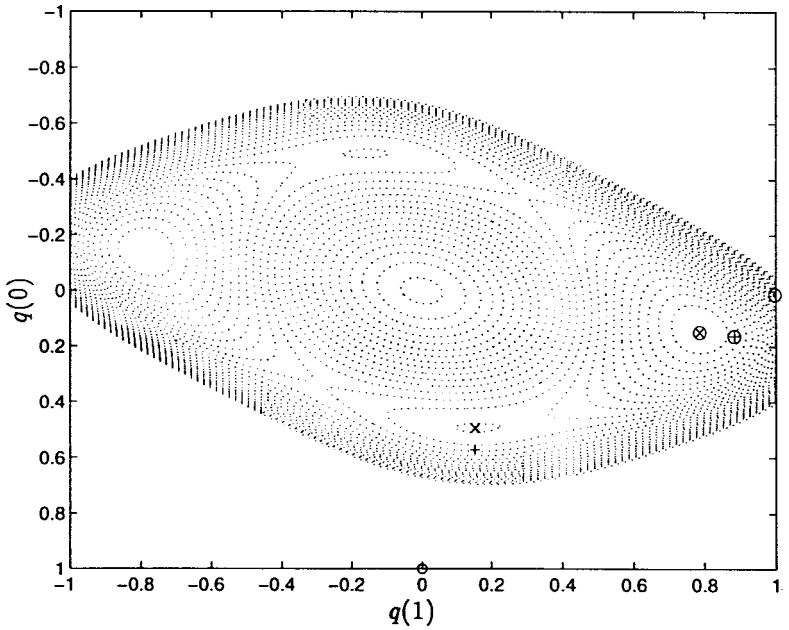

Fig. 3. The contours of the CM cost function. $\times$, $+$, and $\circ$ are the CM, Wiener, and ZF (zero forcing) receivers for $s(k)$ respectively; $\otimes$, $\oplus$, and $\odot$ are the CM, MMSE, and ZF receivers for $s(k-1)$, respectively.

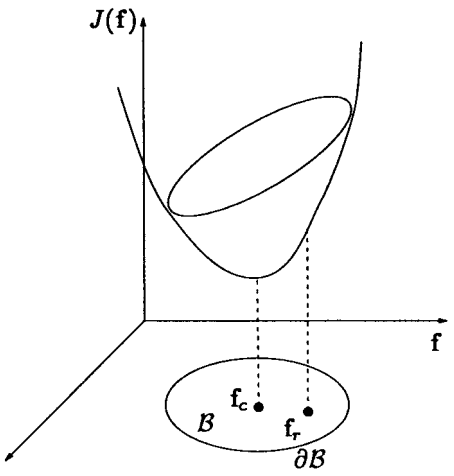

Fig. 4. Illustration of the idea.

interferences, share a common signal model. The approach presented in this paper applies to both cases.

The main results of this paper include

R1: a signal space and matched filter interpretation of CM receivers;

R2: an analytical expression to determine the existence of CM local minima;

R3: an analytical description of regions that contain CM local minima in neighborhoods of Wiener receivers;

R4: upper and lower bounds of the MSE of these CM receivers.

It is important to note that our results apply only to those CM local minima near the Wiener solutions. There are several reasons to study these CM local minima. Due to the quadratic nature of the mean-squared error cost function, any receiver far away from the Wiener receiver has a large MSE. Therefore, if there exist CM local minima in the regions defined in our approach, one of the minima must be the global optimum (in the sense of MSE) CM receiver. The other reason is the strong relationship between the Wiener receiver and the CM receiver, which, to our knowledge, has not been fully exploited except in the well-known noiseless case.

Our results are established based on several assumptions. First, we will restrict our investigation to the real case with the white BPSK source. Extensions to the multilevel signals and the complex case involving QAM type of signals are reported separately [16], [29], [30]. Second, Results R2 to R4 are based on the invertibility of the channel, i.e., the signal can be perfectly recovered. In the application to fractionally spaced equalization, the invertibility condition implies that i) subchannels do not share common zeros and ii) the equalizer is sufficiently long [22]. Although the extension to the case involving noninvertible channels is nontrivial, similar results still hold true [10], [29].

D. Related Work

Existing analysis has been focused on three aspects of CM receiver: i) the existence of local minima; ii) the error surface of CM cost; and iii) the MSE performance of CM receivers. We highlight next some connections among existing results and those presented in this paper. Context-setting surveys can be found in [11], [25], and [31].

Existence of Local Minima in the Absence of Noise: Foschini was the first to show the global convergence of a doubly infinite baud-rate equalizer in the absence of channel noise [7]. For finite or one-sided infinite baud-rate equalizers, the existence of undesirable local minimum has been demonstrated by Ding and his co-workers [2]–[5]. Interestingly, for fractionally spaced finite-impulse response (FIR) equalizers in the absence of noise, the convergence of CMA is global [14] under satisfaction of the associated

identifiability condition [21], [22]. Specifically, in the noise-free case, zero-forcing equalizers are the only local minima. When noise is present, however, the existence of local minima and their locations have not been formally established. Presented in this paper is a result that can be applied directly to fractionally spaced equalizers in *the presence of noise*. We give a sufficient condition for the existence of local minima and specify the region of their locations.

Error Surface of the CM Cost Function: Characteristics of the CM error surface are discussed in [12] for the baud-rate equalizer in the absence of noise. It is shown that the output power of the FIR CM equalizer is between $1/3$ and 1 for BPSK signals, which defines loosely a region for all CM local minima. In this paper, the same bound is obtained for fractionally spaced CM equalizers in the presence of noise.

MSE of CM: In his original paper [8] published in 1980, Godard observed in simulations that the CM equalizer almost achieves the MMSE. Recently, MSE performance of CM has attracted some research interest [15], [20], [23]. All of the approaches taken in [15], [20], and [23] are based on second-order *approximation* of the CM cost function. While such approximations are valid asymptotically as the signal-to-noise ratio (SNR) approaches to infinity, it is not clear that, for a *fixed* finite SNR, the analysis is accurate. Furthermore, the analysis typically presumes the existence of the CM local minimum which has not been formally established. In contrast, the analysis presented in this paper does not involve approximations. When the sufficient condition of the existence of CM local minimum is satisfied, the MSE bounds of a CM receiver are derived. In simulations, our MSE bounds are tighter than those obtained from the second-order approximation approaches.

E. Organization and Notation

This paper is organized as follows. The system model is presented and the problem is formulated in Section II. In Section III, the CM receiver is shown to have a signal space property, which leads to a canonical decomposition of the receiver. In Section IV, locations of the CM local minima are given and MSE bounds of CM receivers are derived. The strong relationship between the CM and Wiener receivers is revealed in both the parameter space and the Hilbert space of the observations. Finally, in Section V, we return to the example in the Introduction to illustrate key ideas of this paper.

The notation used in this paper is standard. Upper and lower case bold letters denote matrices and vectors, respectively. Key symbols are described in the following list.

$(\cdot)^t$	transpose;
$(\cdot)^\dagger$	Moore–Penrose inverse [13, 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;
$\boldsymbol{e}_\nu$	a unit column vector with 1 at the νth entry and zero elsewhere;
$\mathcal{R}^n$	n-dimensional real vector space;

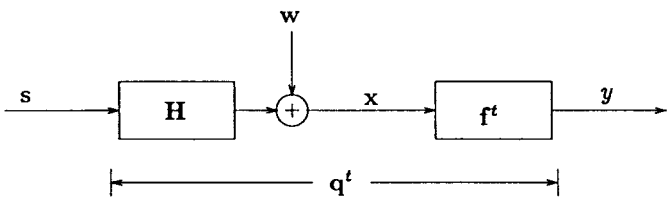

Fig. 5. The model.

$\mathcal{R}^{n\times m}$	the set of all $n \times m$ real matrices;
$\mathcal{R}_{\boldsymbol{A}}$	range of $\boldsymbol{A}\boldsymbol{A}^\dagger$ [13, p. 430];
$\mathcal{R}_{\boldsymbol{A}^\perp}$	range of $\boldsymbol{I}-\boldsymbol{A}\boldsymbol{A}^\dagger$;
$\partial\mathcal{B}$	the boundary of set $\mathcal{B}$;
$\mathcal{A}^c$	the complement of set $\mathcal{A}$.

II. Problem Formulation

Both the FIR equalization of a single user ISI channel and beamforming in array signal processing share the same vector model illustrated in Fig. 5.

$$\boldsymbol{x} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{w} \tag{6}$$

$$y = \boldsymbol{f}^t\boldsymbol{x} = \boldsymbol{q}^t\boldsymbol{s} + \boldsymbol{f}^t\boldsymbol{w} \tag{7}$$

where $\boldsymbol{s} = [s_1, \cdots, s_{n_q}]^t \in \mathcal{R}^{n_q}$ is a vector of the random source signal, $\boldsymbol{w} \in \mathcal{R}^{n_f}$ is the vector of additive noise, $\boldsymbol{H} \in \mathcal{R}^{n_f \times n_q}$ is the (deterministic) channel matrix, $\boldsymbol{x} \in \mathcal{R}^{n_f}$ is the received signal vector. $\boldsymbol{f} \in \mathcal{R}^{n_f}$ is the receiver parameter vector, y is the receiver output, and $\boldsymbol{q} \triangleq \boldsymbol{H}^t\boldsymbol{f} \in \mathcal{R}^{n_q}$ the overall response of combined channel and receiver. The explicit form of the above model for equalization and beamforming can be found in [22].

The following assumptions are made throughout this paper.

- A1: $\boldsymbol{H} \in \mathcal{R}^{n_f \times n_q}$ has full column rank.
- A2: Entries of $\boldsymbol{s}$ are independent and identically distributed (i.i.d.) random variables with equal probability from the set $\{\pm 1\}$.
- A3: Entries of $\boldsymbol{w}$ are i.i.d. Gaussian random variables with zero mean and variance σ^2.

The first assumption satisfies an identifiability condition on the channel [22] and ensures the global convergence of CMA in the absence of the noise [14].

In estimating s_ν, the νth element of vector $\boldsymbol{s}$, we consider the minimization of the following objective functions:

$$\text{CM:} \quad \overline{J}_c(\boldsymbol{f}) \triangleq E\{(|y|^2 - 1)^2\} \tag{8}$$

$$\text{MSE:} \quad \overline{J}_m^\nu(\boldsymbol{f}) \triangleq E\{(y - s_\nu)^2\} \tag{9}$$

$$\text{ZF:} \quad \overline{J}_z^\nu(\boldsymbol{f}) \triangleq \sum_{i\neq\nu,\, q_\nu=1} |q_i|. \tag{10}$$

The CM cost measures the dispersion around the unit circle, the MSE cost measures the dispersion over the signal constellation, and the zero-forcing (ZF) criterion [18] measures the worse case interference.

There are two important differences between the CM and MSE/ZF criteria. First, the Wiener and zero-forcing receivers are nonblind (trained) and can be obtained in closed form. In contrast, there is no closed-form solution for CM receivers.

Our approach is to describe the location of the CM minimum in the neighborhood of the Wiener receiver for a particular ν. Second, the CM cost function is not a function of ν. This of course comes from the nature of a blind receiver. When applied to the equalization problem, ν corresponds to the overall response delay. As shown in the example in the Introduction, CMA may converge to a local minimum near a Wiener receiver with quite a large MSE for a particular ν. However, in nonblind receiver design, an optimal ν can be selected to achieve the least MSE.

In Sections III and IV, we will derive the signal space structure and MSE of CM receivers for $\nu = 1$. For brevity, we shall drop the superscript in $\overline{J}_m^{\nu}(\boldsymbol{f})$ and $\overline{J}_z^{\nu}(\boldsymbol{f})$. Generalization to $\nu \neq 1$ only involves different parameter partition [27], [29].

III. The Signal Space Structures

Our first result shows that both Wiener/ZF and CM receivers are in the "signal subspace" of the observation, which leads to a matched filter interpretation. The so-called signal subspace is defined by $\mathcal{R}_{\boldsymbol{H}}$, the column space of the channel matrix $\boldsymbol{H}$, whereas the noise subspace is its orthogonal complement $\mathcal{R}_{\boldsymbol{H}^{\perp}}$. The signal space property provides the basis of relating the CM receiver to the Wiener receiver in Section IV.

A. The Wiener and Zero-Forcing Receivers and Their Signal Space Properties

For the sake of comparison with the CM receivers, we briefly give the forms of the Wiener and ZF receivers and present their signal space properties.

ZF Receiver: A zero-forcing receiver minimizing $\overline{J}_z(\boldsymbol{f})$, in fact making $\overline{J}_z(\boldsymbol{f}) = 0$ with minimum noise enhancement, is given by

$$\boldsymbol{f}_z = (\boldsymbol{H}^t)^{\dagger}\boldsymbol{e}_1. \tag{11}$$

Since $(\boldsymbol{H}^t)^{\dagger} = \boldsymbol{H}(\boldsymbol{H}^t\boldsymbol{H})^{-1}$ which shares the same column space as $\boldsymbol{H}$, $\boldsymbol{f}_z \in \mathcal{R}_{\boldsymbol{H}}$.

The Wiener Receiver: From (9), we have

$$\overline{J}_m(\boldsymbol{f}) = E\{(\boldsymbol{f}^t\boldsymbol{x} - s_1)^2\} = \boldsymbol{f}^t\boldsymbol{R}\boldsymbol{f} - 2\boldsymbol{e}_1^t\boldsymbol{H}^t\boldsymbol{f} + 1 \tag{12}$$

where

$$\boldsymbol{R} \triangleq E\{\boldsymbol{x}\boldsymbol{x}^t\} = \boldsymbol{H}\boldsymbol{H}^t + \sigma^2\boldsymbol{I}_{n_f}. \tag{13}$$

Minimizing $\overline{J}_m(\boldsymbol{f})$, we have

$$\begin{aligned}\boldsymbol{f}_m &= \boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{e}_1 = \frac{1}{\sigma^2}(\boldsymbol{I}_{n_f} - \boldsymbol{H}(\boldsymbol{H}^t\boldsymbol{H} + \sigma^2\boldsymbol{I}_{n_q})^{-1}\boldsymbol{H}^t)\boldsymbol{H}\boldsymbol{e}_1 \\ &= \boldsymbol{H}\left[\frac{1}{\sigma^2}(\boldsymbol{I}_{n_f} - (\boldsymbol{H}^t\boldsymbol{H} + \sigma^2\boldsymbol{I}_{n_q})^{-1}\boldsymbol{H}^t\boldsymbol{H})\boldsymbol{e}_1\right] \in \mathcal{R}_{\boldsymbol{H}}.\end{aligned} \tag{14}$$

In (14) we have used the matrix inversion lemma.

B. The Signal Space Property of the CM Receivers

The CM receiver minimizes the error between the magnitude of the receiver output and a constant. Using A2 and A3, the cost function of CM can be expressed as (see also [12] for the noiseless case)

$$\begin{aligned}\overline{J}_c(\boldsymbol{f}) &\triangleq E\{(y^2-1)^2\} \\ &= 3||\boldsymbol{f}||_{\boldsymbol{R}}^4 - 2||\boldsymbol{f}||_{\boldsymbol{R}}^2 - 2||\boldsymbol{H}^t\boldsymbol{f}||_4^4 + 1.\end{aligned} \tag{15}$$

Note that $||\boldsymbol{f}||_{\boldsymbol{R}}^2$ is the variance or the power of the receiver output y.

Unlike the Wiener and ZF receivers, there is no closed-form expression for the CM receiver $\boldsymbol{f}_c$ which is a local minimum of $\overline{J}_c(\boldsymbol{f})$. The signal space property of CM receivers must be proved indirectly.

Theorem 1: All local minima of the CM cost function (15) are in the signal subspace $\mathcal{R}_{\boldsymbol{H}}$. Furthermore, the output power of any CM receiver $\boldsymbol{f}_c$ satisfies

$$\frac{1}{3} < ||\boldsymbol{f}_c||_{\boldsymbol{R}}^2 < 1. \tag{16}$$

Proof: See Appendix A.

The signal space results are perhaps not surprising. In [6], Ericson argued that for any "reasonable" criterion of goodness, the optimal receiver includes a matched filter, and therefore, the solution is in the signal subspace. In our case, however, the CM cost function is not "reasonable" in the sense of Ericson when the power of the receiver output is below $1/3$. The signal space result, while intuitively appealing and not unexpected, does not follow directly from Ericson's argument. The important consequence of the signal space result is that the CM receiver is made of a linear combination of filters matched to the columns of the channel matrix $\boldsymbol{H}$.

A tighter upper bound on the power $||\boldsymbol{f}||_{\boldsymbol{R}}$ will be given in Section IV. In the absence of channel noise, the power condition (16) of all CM local minima was given in [12]. Geometrically, this condition implies that *all* CM receivers must be located in an elliptical "shell," illustrated in Fig. 6 for the two-dimensional case. For the ZF receiver $\boldsymbol{f}_z$, according to (11) and (13)

$$||\boldsymbol{f}_z||_{\boldsymbol{R}}^2 = \boldsymbol{e}_1^t\boldsymbol{H}^{\dagger}(\boldsymbol{H}\boldsymbol{H}^t + \sigma^2\boldsymbol{I}_{n_f})(\boldsymbol{H}^t)^{\dagger}\boldsymbol{e}_1 = 1 + \sigma^2||\boldsymbol{f}_z||_2^2.$$

For the Wiener receiver $\boldsymbol{f}_m$, according to (12) and (14)

$$\overline{J}_m(\boldsymbol{f}_m) = 1 - \boldsymbol{f}_m^t\boldsymbol{R}\boldsymbol{f}_m > 0, \Longrightarrow \boldsymbol{f}_m^t\boldsymbol{R}\boldsymbol{f}_m < 1.$$

Therefore, the output power of Wiener receivers is always less than 1, whereas the output power of ZF receivers is greater than 1. As SNR decreases, the output power of a Wiener receiver approaches to zero whereas the output power of a CM receiver is always above $1/3$. This condition, particularly the lower bound, is useful in determining if there is a CM local minima near the Wiener receiver.

C. A Canonical Receiver Structure

From the signal space structure of the receiver, we derive a useful canonical decomposition and give the equivalent MSE and CM costs as functions of the overall impulse response $\boldsymbol{q}$ defined in (6) and (7).

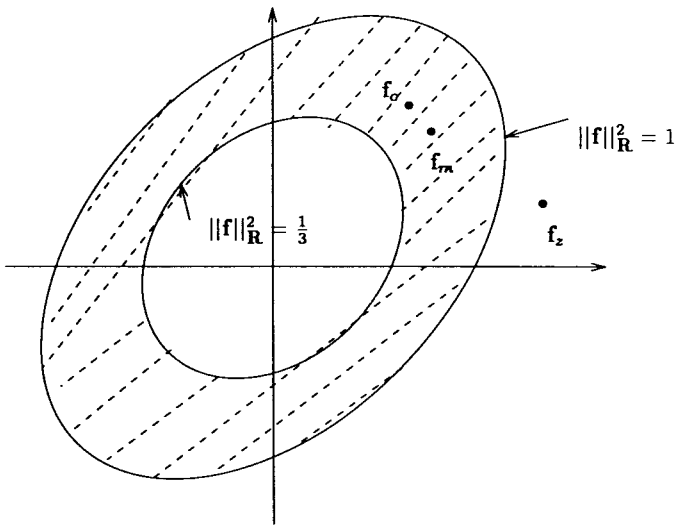

Fig. 6. Region of CM local minima.

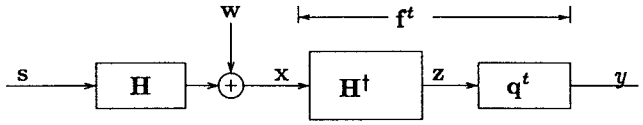

Fig. 7. The decomposition of the ZF receivers $\boldsymbol{H}^\dagger$ and the linear combiner $\boldsymbol{q}$.

The Zero-Forcing Front-End: Consider the combined channel $\boldsymbol{q} = \boldsymbol{H}^t\boldsymbol{f}$. If the receiver $\boldsymbol{f}$ is in the signal space, then

$$\boldsymbol{f} = (\boldsymbol{H}^t)^\dagger \boldsymbol{q} \tag{17}$$

i.e., $\boldsymbol{f}$ is made of the zero-forcing receiver bank at the front-end followed by a "linear combiner" $\boldsymbol{q}$ which turns out to be the overall impulse response. This leads to the structure of the receiver in Fig. 7. Define $\boldsymbol{v}$ as the noise output of the ZF receivers, i.e., $\boldsymbol{v} = \boldsymbol{H}^\dagger \boldsymbol{w}$, which is colored noise with covariance of $\sigma^2(\boldsymbol{H}^t\boldsymbol{H})^{-1}$. The model in Fig. 7 is now equivalent to that in Fig. 8.

While it may not be practical for implementation because of the noise enhancement problem of the ZF front-end, the canonical structure given in Fig. 8 offers an important decomposition at the conceptual level. The design of the receiver can be considered, without loss of generality, as a two-step process that i) eliminates the interference and ii) compensates the colored noise $\boldsymbol{v}$ by processing the output $\boldsymbol{z}$ of the ZF front-end through the linear combiner $\boldsymbol{q}$. It is the latter step where all receivers with the signal space property differ.

In the case of Wiener receiver, it turns out that the design of linear combiner $\boldsymbol{q}$ depends only on the covariance matrix of $\boldsymbol{z}$, the output of the ZF front-end, given by

$$\begin{aligned}\boldsymbol{\Phi} &\triangleq E\{\boldsymbol{z}\boldsymbol{z}^t\} = E\{\boldsymbol{s}\boldsymbol{s}^t\} + E\{\boldsymbol{v}\boldsymbol{v}^t\} \\ &= \boldsymbol{I}_{n_q} + \sigma^2(\boldsymbol{H}^t\boldsymbol{H})^{-1} = \boldsymbol{H}^\dagger \boldsymbol{R}(\boldsymbol{H}^t)^\dagger.\end{aligned} \tag{18}$$

Cost Functions in the Space of Linear Combiners: The canonical structure of receivers in the signal space enables us to carry out the analysis in the parameter space of the linear combiner $\boldsymbol{q}$, the overall response of combined channel and receiver. The CM, Wiener, and ZF receivers are different in their choices of the combiner $\boldsymbol{q}$. In the sequel, we shall use the equivalent MSE and CM costs as functions of the linear combiner

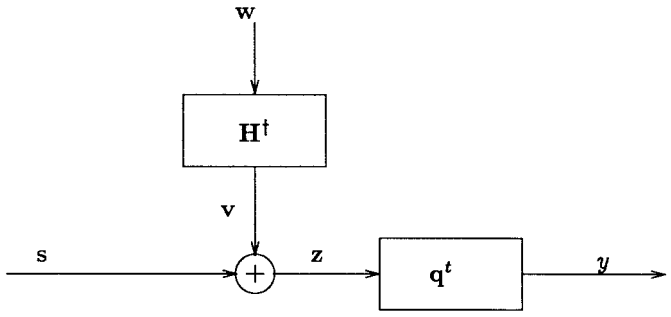

Fig. 8. The canonical decomposition.

$$\text{MSE:} \quad J_m(\boldsymbol{q}) \triangleq \overline{J}_m((\boldsymbol{H}^t)^\dagger \boldsymbol{q}) = \boldsymbol{q}^t\boldsymbol{\Phi}\boldsymbol{q} - 2\boldsymbol{q}^t\boldsymbol{e}_1 + 1 \tag{19}$$

$$\text{CM:} \quad J(\boldsymbol{q}) \triangleq \overline{J}_c((\boldsymbol{H}^t)^\dagger \boldsymbol{q}) = 3||\boldsymbol{q}||^4_{\boldsymbol{\Phi}} - 2||\boldsymbol{q}||^2_{\boldsymbol{\Phi}} - 2||\boldsymbol{q}||^4_4 + 1 \tag{20}$$

where $\boldsymbol{\Phi}$ is defined in (18). Note that the output power of CM receivers can be obtained from the norm of receiver vector $\boldsymbol{f}$ and $\boldsymbol{q}$, i.e., $E\{|y|^2\} = ||\boldsymbol{f}||^2_{\boldsymbol{R}} = ||\boldsymbol{q}||^2_{\boldsymbol{\Phi}}$.

IV. Wiener and CM Receivers

Most of our results rest on the connection between the Wiener and CM receivers. Of particular importance is a partition of the parameter based on the concept of a (conditional) *unbiased estimator*. The region containing the CM receiver is then determined only by the bias and interference of the Wiener receivers. The interference can be either the intersymbol interference in a single-user equalizer or the interuser interference in a multiuser beamformer.

A. The Conditionally Unbiased Estimator

Consider the overall impulse response $\boldsymbol{q}$ of the combined channel and receiver in Fig. 5. From (7), the estimate of signal s_1 is given by

$$y = \boldsymbol{q}^t\boldsymbol{s} + \boldsymbol{f}^t\boldsymbol{w} = q_1 s_1 + \sum_{i=2}^{n_q} q_i s_i + \boldsymbol{f}^t\boldsymbol{w} \tag{21}$$

where q_1 represents the gain of the signal of interest s_1. The other components $\{q_2, \cdots, q_{n_q}\}$ are gains of the interference. By parameterizing $\boldsymbol{q}$ as

$$\boldsymbol{q} \triangleq \theta\begin{pmatrix}1\\ \boldsymbol{q}_{\mathrm{I}}\end{pmatrix} \tag{22}$$

$||\boldsymbol{q}_{\mathrm{I}}||$ measures the intersymbol/user interference (ISI/IUI) of the receiver [18, p. 541]. This parameterization results in the realization of $\boldsymbol{q}$ in Fig. 9. The receiver output y is a scaled version of the unbiased estimate u of s_1 (conditioned on s_1), i.e., $y = \theta u$. The significance of using u for detection has recently been emphasized by Cioffi *et al.* [1].

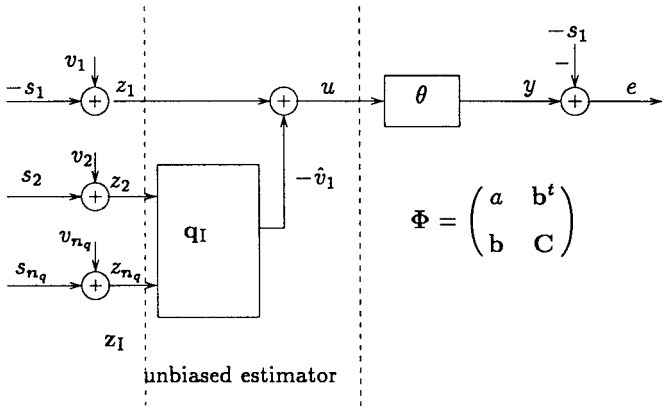

Fig. 9. Decomposition into unbiased estimator and gain.

Let $\boldsymbol{z}$ and $\boldsymbol{\Phi}$ be partitioned according to (22), i.e.,

$$\boldsymbol{z} \triangleq \begin{pmatrix} z_1 \\ \boldsymbol{z}_\mathrm{I} \end{pmatrix}, \quad \boldsymbol{\Phi} = \begin{pmatrix} E\{z_1 z_1\} & E\{z_1 \boldsymbol{z}_\mathrm{I}^t\} \\ E\{\boldsymbol{z}_\mathrm{I} z_1\} & E\{\boldsymbol{z}_\mathrm{I} \boldsymbol{z}_\mathrm{I}^t\} \end{pmatrix} \triangleq \begin{pmatrix} a & \boldsymbol{b}^t \\ \boldsymbol{b} & \boldsymbol{C} \end{pmatrix} \tag{23}$$

with appropriate definitions of a, $\boldsymbol{b}$, and $\boldsymbol{C}$. As shown in Fig. 9, $\boldsymbol{q}_\mathrm{I}$ is the part of the combiner that reduces the noise v_1 using $\boldsymbol{z}_\mathrm{I}$. It also introduces the interference due to s_i, $i = 2, \cdots, n_q$. $|1-\theta|$ is the magnitude of the receiver bias conditional on s_1, i.e., $|1 - E\{u|s_1\}| = |1-\theta|$. The parameterization in (22) that decomposes $\boldsymbol{q}$ into the unbiased estimator and a gain factor θ holds the key to connecting CM receivers with Wiener receivers. This will be further discussed after Lemma 1.

B. The Wiener Receiver

The Wiener receiver $\boldsymbol{q}_m$ for s_1, obtained by minimizing $J_m(\boldsymbol{q})$ in (20), is given by

$$\boldsymbol{q}_m = \boldsymbol{\Phi}^{-1} \boldsymbol{e}_1. \tag{24}$$

Under the decomposition in the preceding section, we have the following properties of the Wiener receiver. The proof of Lemma 1 is important in establishing the geometrical description in the Hilbert space of the observations.

Lemma 1: Let $\boldsymbol{q}_m$, $\boldsymbol{q}_{m\mathrm{I}}$, θ_m, u_m, y_m, and e_m be the corresponding terms in Fig. 9 for the Wiener receiver. We then have

a) $\boldsymbol{q}_{m\mathrm{I}}$ and θ_m are given by

$$\boldsymbol{q}_m = \theta_m \begin{pmatrix} 1 \\ \boldsymbol{q}_{m\mathrm{I}} \end{pmatrix} = \underbrace{\frac{1}{a - \boldsymbol{b}^t \boldsymbol{C}^{-1} \boldsymbol{b}}}_{\theta_m} \begin{pmatrix} 1 \\ -\boldsymbol{C}^{-1} \boldsymbol{b} \end{pmatrix}; \tag{25}$$

b) the MMSE, bias and output power of the Wiener receiver are related by

$$\theta_m = 1 - J_m(\boldsymbol{q}_m) = \|\boldsymbol{q}_m\|_{\boldsymbol{\Phi}}^2 \le 1; \tag{26}$$

c) for any receiver $\boldsymbol{q} = \theta \begin{pmatrix} 1 \\ \boldsymbol{q}_\mathrm{I} \end{pmatrix}$, let δ^2 be the extra mean-squared error of the conditionally unbiased estimate u over the conditionally unbiased MMSE estimate u_m, i.e.,

$$\delta^2 \triangleq E\{|u - s_1|^2\} - E\{|u_m - s_1|^2\} \tag{27}$$

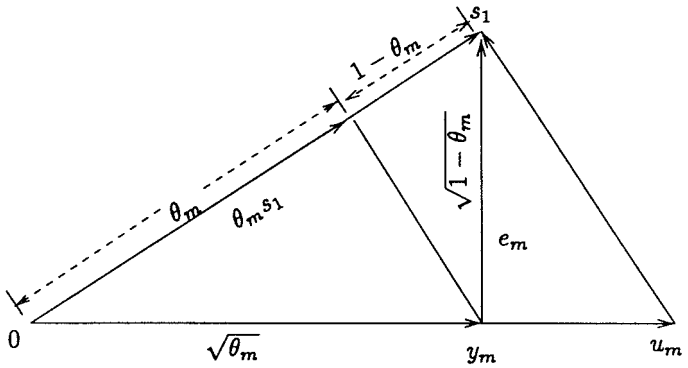

Fig. 10. Wiener receiver in the Hilbert space of the observations.

then

$$\delta^2 = E\{|u - u_m|^2\} = \|\boldsymbol{q}_I - \boldsymbol{q}_{m\mathrm{I}}\|_{\boldsymbol{C}} \tag{28}$$

$$E\{y^2\} = \|\boldsymbol{q}\|_{\boldsymbol{\Phi}}^2 = \theta^2 \left(\frac{1}{\theta_m} + \delta^2 \right); \tag{29}$$

d) for any receiver $\boldsymbol{q}$ with output power less than that of the Wiener receiver, i.e., $\|\boldsymbol{q}\|_{\boldsymbol{\Phi}}^2 \le \|\boldsymbol{q}_m\|_{\boldsymbol{\Phi}}^2$, the bias of $\boldsymbol{q}$ is greater than that of the Wiener receiver, i.e., $\theta \le \theta_m$.

Proof: Although the proof can be obtained algebraically, we present a "Pythagorean" proof consistent with the geometrical approach in this paper. Relationships of various random variables in their Hilbert space are depicted in Figs. 10 and 11.

a) Note that $\boldsymbol{C} = E\{\boldsymbol{z}_\mathrm{I} \boldsymbol{z}_\mathrm{I}^t\}$ is the covariance of the interference $\boldsymbol{z}_\mathrm{I} = [z_2, \cdots, z_{n_q}]^t$, $\boldsymbol{b}$ is the crosscorrelation between the interference $\boldsymbol{z}_\mathrm{I}$ and v_1, i.e., $\boldsymbol{b} = E\{\boldsymbol{z}_\mathrm{I} v_1\}$. The MMSE noise cancellation of v_1 using $\boldsymbol{z}_\mathrm{I}$ is then given by $-\boldsymbol{C}^{-1}\boldsymbol{b}$ which leads to the unbiased Wiener receiver $[1, \boldsymbol{q}_{m\mathrm{I}}^t]^t$. The variance of the unbiased Wiener receiver is $a - \boldsymbol{b}^t \boldsymbol{C}^{-1} \boldsymbol{b}$; therefore, the scaling factor of the Wiener receiver is given by $1/(a - \boldsymbol{b}^t \boldsymbol{C}^{-1} \boldsymbol{b})$.

b) From the principle of orthogonality, y_m is orthogonal to the error e_m, as shown in Fig. 10. The projection of y_m on the s_1 is $\theta_m s_1$. From Fig. 10, due to the triangle similarity

$$\frac{E\{|y_m|^2\}}{E\{|s_1|^2\}} = \frac{E\{|\theta_m s_1|^2\}}{E\{|y_m|^2\}} \tag{30}$$

where $E\{|y_m|^2\} = \|\boldsymbol{q}\|_{\boldsymbol{\Phi}}^2$ and $E\{|s_1|^2\} = 1$. We then have $E\{|y_m|^2\} = \theta_m$. Similarly, we can prove that

$$J_m(\boldsymbol{q}_m) = E\{|e_m|^2\} = 1 - \theta_m.$$

c) Consider the relationship between the output y of any receiver $\boldsymbol{q}$ and y_m shown in Fig. 11. By scaling y and y_m, we have the corresponding unbiased estimators u and u_m, respectively. Since the orthogonal projections of u and u_m on s_1 are both s_1, we have $u - u_m \perp u_m$. Hence

$$\begin{aligned} \delta^2 &= E\{|u - s_1|^2\} - E\{|u_m - s_1|^2\} \\ &= E\{|u - u_m|^2\} = \|\boldsymbol{q}_I - \boldsymbol{q}_{m\mathrm{I}}\|_{\boldsymbol{C}}. \end{aligned} \tag{31}$$

Furthermore, from Fig. 11, we have

$$E\{|y|^2\} = \theta^2 E\{|u|^2\} = \theta^2 (E\{|u_m|^2\} + E\{|u - u_m|^2\}) \tag{32}$$

$$= \theta^2 \left(\frac{E\{|y_m|^2\}}{\theta_m^2} + \delta^2 \right) = \theta^2 \left(\frac{1}{\theta_m} + \delta^2 \right). \tag{33}$$

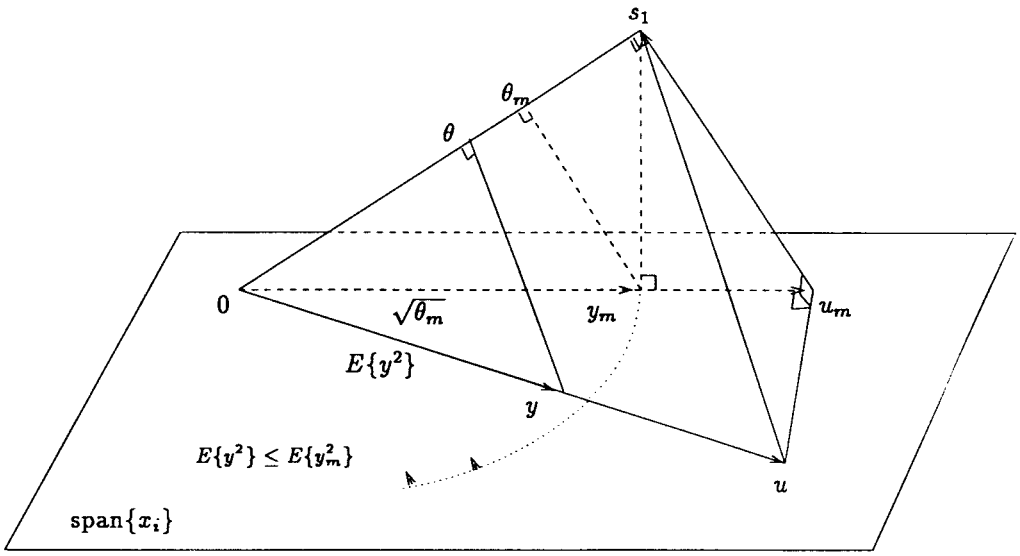

Fig. 11. The Hilbert space of the observations.

d) From Fig. 11, this is immediate. Mathematically, from (29), when $E\{|y|^2\} \le E\{|y_m|^2\}$

$$\theta^2 = \frac{E\{|y|^2\}}{\dfrac{1}{\theta_m} + \delta^2} \le \frac{\theta_m^2}{1+\theta_m\delta^2} \le \theta_m^2. \tag{34}$$

□

Remarks: From (29), the power of a receiver output can be represented by θ and δ, where $|1-\theta|$ and δ^2 are the magnitude of the bias and the extra mean-squared error of the conditionally unbiased estimator. The CM cost function (20) is then given by

$$J(\boldsymbol{q}) = 3\left(\delta^2 + \frac{1}{\theta_m}\right)^2\theta^4 - 2\left(\delta^2 + \frac{1}{\theta_m}\right)\theta^2 - 2\theta^4\left(1 + ||\boldsymbol{q}_{\mathrm{I}}||_4^4\right) + 1. \tag{35}$$

In the next section, $||\boldsymbol{q}_{\mathrm{I}}||_4$ in (35) can be bounded in terms of δ. Thus the cost function is reduced to a function of two important variables: θ and δ, both of which are direct consequence of the parameterization given in (22).

C. The Location of CMA Local Minima

The basic idea, as discussed in the Introduction, is to find a region $\mathcal{B}$ in the linear combiner space such that all points on its boundary have costs greater than the cost of (at least) one point in this region in Fig. 4. Consequently, there is a local minimum of CM cost function in $\mathcal{B}$. In our approach, the selection of $\boldsymbol{q}_r$ and $\mathcal{B}$ is based on the structure of the Wiener receiver. We shall also give the corresponding region in the Hilbert space of the observations.

The Reference $\boldsymbol{q}_r$: In relating the CM with the Wiener receiver, we choose, in the direction of the Wiener solution $\boldsymbol{q}_m$, the reference point $\boldsymbol{q}_r \triangleq \theta_r\binom{1}{\boldsymbol{q}_{m\mathrm{I}}}$ with the minimum CM cost. From (35), θ_r is obtained by minimizing the following cost function with respect to θ:

$$J\left(\theta\binom{1}{\boldsymbol{q}_{m\mathrm{I}}}\right) = \frac{3}{\theta_m^2}\theta^4 - \frac{2}{\theta_m}\theta^2 - 2(1+||\boldsymbol{q}_{m\mathrm{I}}||_4^4)\theta^4 + 1. \tag{36}$$

Setting $\partial J/\partial\theta = 0$ leads to

$$\theta_r^2 = \frac{\theta_m}{3 - 2\theta_m^2 - 2\theta_m^2||\boldsymbol{q}_{m\mathrm{I}}||_4^4} \tag{37}$$

$$J(\boldsymbol{q}_r) = 1 - \frac{1}{3 - 2\theta_m^2 - 2\theta_m^2||\boldsymbol{q}_{m\mathrm{I}}||_4^4} = 1 - \frac{\theta_r^2}{\theta_m}. \tag{38}$$

Note that $\theta_r < \theta_m$ when $\theta_m > 2/3$. This can be verified numerically by using

$$||\boldsymbol{q}_{m\mathrm{I}}||_4^2 \le ||\boldsymbol{q}_{m\mathrm{I}}||_2^2 = \frac{||\theta_m\boldsymbol{q}_{m\mathrm{I}}||^2}{\theta_m^2} < \frac{J_m(\boldsymbol{q}_m) - (1-\theta_m)^2}{\theta_m^2} = \frac{1-\theta_m}{\theta_m}. \tag{39}$$

The Region $\mathcal{B}$ in the Parameter Space: In the parameter space of $\boldsymbol{q}$, the definition of the region $\mathcal{B}$ is illustrated in Fig. 12. Consider the cone $\mathcal{K}_p(\boldsymbol{q}_m, \delta)$ with $\boldsymbol{q}_m$ as its axis

$$\mathcal{K}_p(\boldsymbol{q}_m, \delta) \triangleq \left\{\boldsymbol{q} = \theta\binom{1}{\boldsymbol{q}_I} : ||\boldsymbol{q}_I - \boldsymbol{q}_{m\mathrm{I}}||_{\boldsymbol{C}} \le \delta\right\}. \tag{40}$$

The "radius" δ, defined in Lemma 1, is also the radius of the $\mathcal{K}_p(\boldsymbol{q}_m, \delta)$ sliced at $\theta = 1$ corresponding to the zero-forcing solution. The region $\mathcal{B}$ considered here is a subset of $\mathcal{K}_p(\boldsymbol{q}_m, \delta)$ by slicing $\mathcal{K}_p(\boldsymbol{q}_m, \delta)$ at θ_L and θ_U. Specifically

$$\mathcal{B}_p(\boldsymbol{q}_m, \delta, \theta_L, \theta_U) = \left\{\boldsymbol{q} \triangleq \theta\binom{1}{\boldsymbol{q}_I} : \theta_L \le \theta \le \theta_U, \; ||\boldsymbol{q}_I - \boldsymbol{q}_{m\mathrm{I}}||_{\boldsymbol{C}} \le \delta\right\}. \tag{41}$$

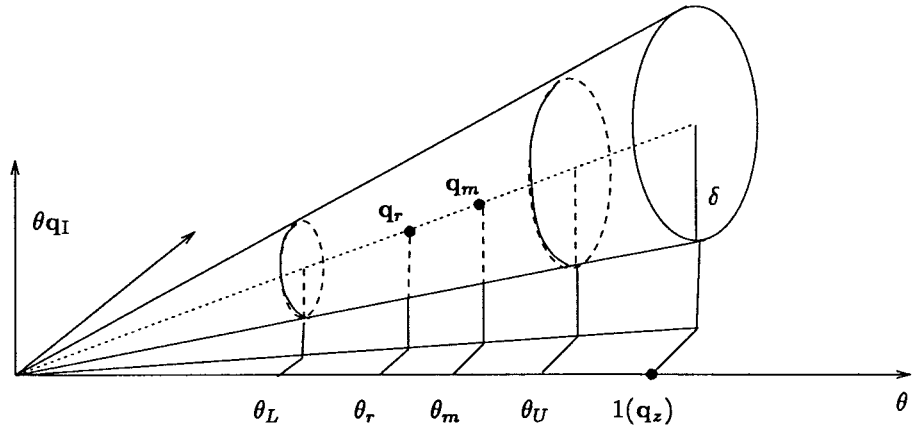

Fig. 12. The region $\mathcal{B}$ in the parameter space $\mathcal{R}^{nq}$.

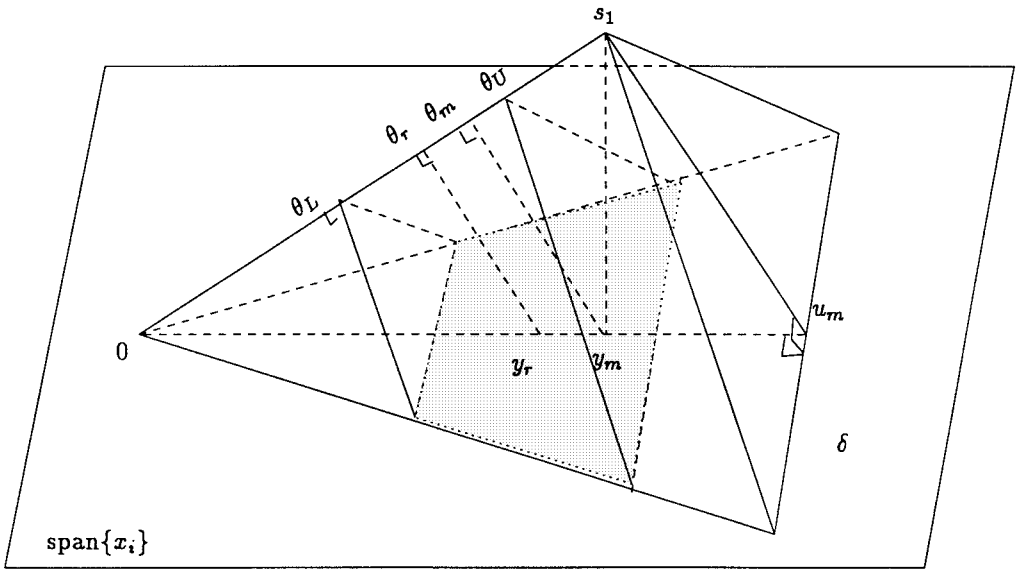

Fig. 13. The region $\mathcal{B}$ in the Hilbert space of the observations.

The Region $\mathcal{B}$ in the Hilbert Space of the Observations: The corresponding description of $\mathcal{B}_p(\boldsymbol{q}_m, \delta, \theta_L, \theta_U)$ in the Hilbert space of the observations is shown in Fig. 13. The cone $\mathcal{K}_p(\boldsymbol{q}_m, \delta)$ and the region $\mathcal{B}_p(\boldsymbol{q}_m, \delta, \theta_L, \theta_U)$ defined in (40) and (41) become

$$\mathcal{K}_o(y_m, \delta) \triangleq \left\{ y\colon E\left\{ \left(\frac{1}{\theta} y - \frac{1}{\theta_m} y_m \right)^2 \right\} \le \delta^2 \right\} \tag{42}$$

$$\mathcal{B}_o(y_m, \delta, \theta_L, \theta_U) \triangleq \{ y \in \mathcal{K}_o(y_m, \delta)\colon \theta_L \le \mathcal{P}_{s_1}(y) \le \theta_U \} \tag{43}$$

where $\mathcal{P}_{s_1}(y)$ is the orthogonal projection of y in the direction of s_1. The radius δ is the extra MSE between the unbiased Wiener receiver and the unbiased receiver $[1, \boldsymbol{q}_{\mathrm{I}}^t]^t$. The relationships among the bias, MSE and the output power are shown in Fig. 13. Comparing with Fig. 12, the sliced cone $\mathcal{B}(\theta_L, \theta_U, \delta_U)$ corresponds to the shaded trapezoid in Fig. 13.

One of our objectives is to quantify θ_U and θ_L. Our goal now is to find as small a δ and as small a $\theta_U - \theta_L$ as possible such that all the points on the surface of $\mathcal{B}_p(\boldsymbol{q}_m, \delta, \theta_L, \theta_U)$ have costs greater than the cost at $\boldsymbol{q}_r$. Evaluating all points on the boundary $\partial\mathcal{B}$ is difficult. The following lemma allows us to bound the CM cost of $\boldsymbol{q}$ on $\partial\mathcal{K}_p(\boldsymbol{q}_m, \delta)$ using the bias and the intersymbol/user interference.

Lemma 2: Given the Wiener receiver $\boldsymbol{q}_m = \theta_m \binom{1}{\boldsymbol{q}_{m\mathrm{I}}}$. Let $\boldsymbol{q}_r = \theta_r \binom{1}{\boldsymbol{q}_{m\mathrm{I}}}$ be the reference defined in (37), (38). Then, for any $\boldsymbol{q} = \theta \binom{1}{\boldsymbol{q}_{\mathrm{I}}} \in \partial\mathcal{K}_p(\boldsymbol{q}_m, \delta)$

$$J(\boldsymbol{q}) - J(\boldsymbol{q}_r) \ge c_2(\delta)\theta^4 + c_1(\delta)\theta^2 + c_0 \quad (\text{equality holds iff } \delta = 0) \tag{44}$$

where

$$c_0 = \frac{1}{3 - 2\theta_m^2 - 2\theta_m^2 ||\boldsymbol{q}_{m\mathrm{I}}||_4^4} \tag{45}$$

$$c_1(\delta) = -2\left(\delta^2 + \frac{1}{\theta_m} \right) \tag{46}$$

$$c_2(\delta) = 3\left(\delta^2 + \frac{1}{\theta_m} \right)^2 - 2(1 + (\delta + ||\boldsymbol{q}_{m\mathrm{I}}||_4)^4). \tag{47}$$

Proof: See Appendix B.

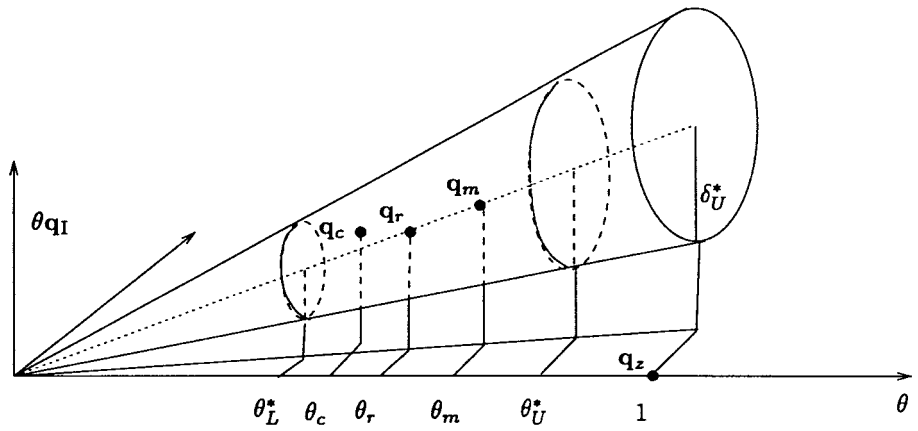

Fig. 14. The CM receiver output $\boldsymbol{q}_c$ in the parameter space $\mathcal{R}^{nq}$.

From this lemma, we can see that, on $\partial\mathcal{K}_p(\boldsymbol{q}_m, \delta)$, 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 $\mathcal{B}_p(\boldsymbol{q}_m, \delta, \theta_L, \theta_U)$ is obtained by choosing θ_L, θ_U, and δ such that $J(\boldsymbol{q}) - J(\boldsymbol{q}_r) > 0$ for all $\boldsymbol{q} \in \partial\mathcal{B}$. If such θ_L, θ_U, and δ exist, we have located at least one CM local minimum.

Theorem 2: Consider a slice of a cone $\mathcal{B}_p(\boldsymbol{q}_m, \delta, \theta_L, \theta_U)$ defined in (41), and define

$$D(\delta) \triangleq c_1(\delta)^2 - 4c_2(\delta)c_0 \tag{48}$$

where $c_2(\delta)$, $c_1(\delta)$, and c_0 are given in (45)–(47). Under the condition that $J_m(\boldsymbol{q}_m) < 1/3$, if $D(||\boldsymbol{q}_{m\mathrm{I}}||_2) < 0$, then there exists a local minimum in $\mathcal{B}_p(\boldsymbol{q}_m, \delta_U^*, \theta_L^*, \theta_U^*)$ where δ_U^* is the smallest positive root of $D(\delta)$

$$\theta_L^* = \min_{0\le\delta\le\delta_U^*} \sqrt{\frac{-c_1(\delta) - \sqrt{c_1(\delta)^2 - 4c_2(\delta)c_0}}{2c_2(\delta)}} \tag{49}$$

$$\theta_U^* = \max_{0\le\delta\le\delta_U^*} \sqrt{\frac{-c_1(\delta) + \sqrt{c_1(\delta)^2 - 4c_2(\delta)c_0}}{2c_2(\delta)}}. \tag{50}$$

Furthermore,

$$\lim_{\sigma\to 0} \mathcal{B}_p(\boldsymbol{q}_m, \delta_U^*, \theta_L^*, \theta_U^*) = \lim_{\sigma\to 0}\{\boldsymbol{q}_c\} = \lim_{\sigma\to 0}\{\boldsymbol{q}_m\} = \{\boldsymbol{q}_z\}. \tag{51}$$

Proof: See Appendix C.

Remarks:

- Given $\boldsymbol{q}_m = \theta_m[1, \boldsymbol{q}_{m\mathrm{I}}^t]^t$, this theorem can be used i) to detect the existence of a CM local minimum by checking the sufficient condition $D(||\boldsymbol{q}_{m\mathrm{I}}||_2) < 0$ and ii) to determine the region $\mathcal{B}_p(\boldsymbol{q}_m, \delta_U^*, \theta_L^*, \theta_U^*)$.
- The condition in the theorem is *sufficient* and can be relaxed. In [29], it is shown that if $c_2(\delta) > 0$ for $0 \le \delta \le \delta_U^*$, then there exists a local minimum in $\mathcal{B}_p(\boldsymbol{q}_m, \delta_U^*, \theta_L^*, \theta_U^*)$.
- Equation (51) shows that the region $\mathcal{B}_p(\boldsymbol{q}_m, \delta_U^*, \theta_L^*, \theta_U^*)$ shrinks to the Wiener/ZF receiver as noise vanishes.

D. The Bias, ISI/IUI, and Output Power of CM Receivers

Further characterization of CM local minima is possible once the region $\mathcal{B}_p(\boldsymbol{q}_m, \delta_U^*, \theta_L^*, \theta_U^*)$ containing a CMA local minimum is obtained. Of particular interest are the bias, the intersymbol/user interference, and the output power of the CM receiver. The next result can be used to further tighten the region $\mathcal{B}_p(\boldsymbol{q}_m, \delta_U^*, \theta_L^*, \theta_U^*)$.

Theorem 3: Let $\boldsymbol{q}_m = \theta_m \begin{pmatrix} 1 \\ \boldsymbol{q}_{m\mathrm{I}} \end{pmatrix}$ be the Wiener receiver. a) Suppose

$$\boldsymbol{q}_c = \theta_c[1, \boldsymbol{q}_{c\mathrm{I}}^t]^t \in \mathcal{B}_p(\boldsymbol{q}_m, \delta_U^*, \theta_L^*, \theta_U^*)$$

is a CM minimum. If $\theta_m > 0.6$, or equivalently by (26), $J_m(\boldsymbol{q}_m) < 0.4$, the output power of CM receiver is less than that of the Wiener receiver, i.e., $1/3 < ||\boldsymbol{q}_c||_\Phi^2 < ||\boldsymbol{q}_m||_\Phi^2 < 1$, and the bias of the CM receiver is greater than that of the Wiener receiver, i.e., $\theta_c \le \theta_m$. b) Suppose $\boldsymbol{q}_c$ is the CM minimum in $\mathcal{B}_p(\boldsymbol{q}_m, \delta_U^*, \theta_L^*, \theta_U^*)$, then $||\boldsymbol{q}_{c\mathrm{I}}||_4 \ge ||\boldsymbol{q}_{m\mathrm{I}}||_4$.

Proof: See Appendix D.

Summary of the Relationships: We now summarize the relationship among the CM, Wiener, and ZF receivers developed so far. Fig. 14 shows the region $\mathcal{B}_p(\boldsymbol{q}_m, \delta_U^*, \theta_L^*, \theta_U^*)$ and the receivers in the parameter space. We note that $\theta_r \le \theta_m$, and in general, the CM receiver is closer to the Wiener receiver than the ZF receiver. Fig. 15 shows the relationship between the outputs of the CM and the Wiener receivers and their corresponding unbiased estimators. Again, the output power of the CM receivers output must be less than that of the Wiener receiver (in the shaded region). Hence, we can further obtain a tighter θ_U^* via

$$\theta_U^* \triangleq \min\left\{\theta_m, \max_{0\le\delta\le\delta_U^*} \sqrt{\frac{-c_1(\delta) + \sqrt{c_1(\delta)^2 - 4c_2(\delta)c_0}}{2c_2(\delta)}}\right\}. \tag{52}$$

E. MSE of CM Receivers

One of the most important properties of a CM receiver is its MSE performance. The result of Theorem 2 enables us to give performance bounds on MSE.

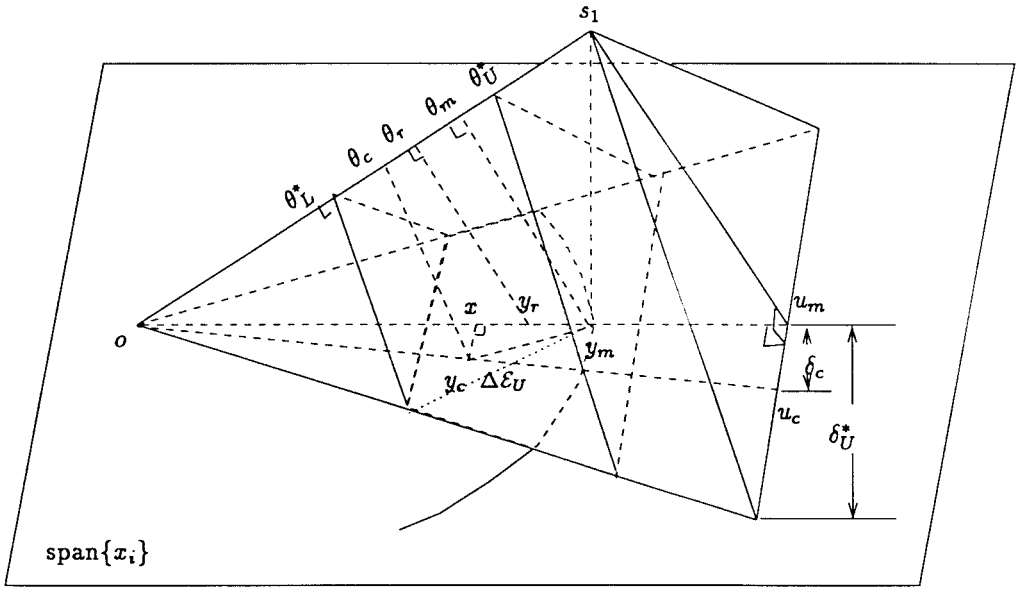

Fig. 15. The CM receiver output y_c in the Hilbert space of the observations.

Theorem 4: Suppose $J_m(\boldsymbol{q}_m) < 1/3$. Let $\Delta\mathcal{E}$ be the extra MSE of a CM minimum $\boldsymbol{q}_c \in \mathcal{B}_p(\boldsymbol{q}_m, \delta_U^*, \theta_L^*, \theta_U^*)$ versus the MSE of the Wiener receiver, i.e., $\Delta\mathcal{E} \triangleq J_m(\boldsymbol{q}_c) - J_m(\boldsymbol{q}_m)$. Then

$$\underbrace{\frac{(\theta_U^* - \theta_m)^2}{\theta_m}}_{\Delta\mathcal{E}_L} \le \Delta\mathcal{E} \le \underbrace{\frac{(\theta_L^* - \theta_m)^2}{\theta_m} + (\theta_U^* \delta_U^*)^2}_{\Delta\mathcal{E}_U}. \tag{53}$$

Let $\widehat{\Delta\mathcal{E}}$ be the extra MSE of the reference, i.e., $\widehat{\Delta\mathcal{E}} \triangleq J_m(\boldsymbol{q}_r) - J_m(\boldsymbol{q}_m)$. Then

$$\widehat{\Delta\mathcal{E}} = \frac{(\theta_r - \theta_m)^2}{\theta_m} = \left(\frac{1}{\sqrt{3 - 2\theta_m^2(1 + \|\boldsymbol{q}_{m\mathrm{I}}\|_4^4)}} - \sqrt{\theta_m} \right)^2 \tag{54}$$

$$= \frac{9}{4} J_m^2(\boldsymbol{q}_m) + O(J_m^2(\boldsymbol{q}_m)). \tag{55}$$

Proof: See Appendix E.

Both $\Delta\mathcal{E}_U$ and $\widehat{\Delta\mathcal{E}}$ can be seen directly from the Hilbert space in Fig. 15. $\Delta\mathcal{E}_U$ is related to the longest distance between the y_m and any point in the shaded region. $\widehat{\Delta\mathcal{E}}$ is the distance between y_r and y_m.

The bounds involve the computation of θ_L^*, θ_U^*, δ_U^*. If the size of the $\mathcal{B}$ is small, the MSE of the CM receiver can be approximated by the MSE at the reference, and the CM receiver is approximated by the reference point which is the scaled version of the Wiener receiver. Interestingly, the MSE based on the quadratic approximations in [15] and [20] has the same second-order term $\frac{9}{4} J_m^2(\boldsymbol{q}_m)$ as $\widehat{\Delta\mathcal{E}}$. It is, however, important to note that $\widehat{\Delta\mathcal{E}}$ in (54) is different from the approximation and is shown to be more accurate in simulations.

V. An Example

In this section, we return to the example given in the Introduction to illustrate the key ideas in this paper. While this example is artificial, a more elaborate study of practical multipath channels can be found in [28] and [29].

The channel impulse response is $\{1, 1, 3, 0.5\}$ and the equalizer has two taps. The corresponding channel matrix is a 2×2 matrix given by (see [28] for the definition of $\boldsymbol{H}$)

$$\boldsymbol{H} = \begin{pmatrix} 1 & 3 \\ 1 & 0.5 \end{pmatrix}. \tag{56}$$

The receiver $\boldsymbol{f}$ and the overall impulse response vector $\boldsymbol{q}$ are two-dimensional vectors, i.e., $n_f = 2$, $n_q = 2$. The signal-to-noise ratio (SNR) is defined by

$$\mathrm{SNR} \triangleq 10 \log_{10} \left(\frac{\|\boldsymbol{H}\|_F^2}{2\sigma^2} \right) \tag{57}$$

where $\|\cdot\|_F$ denotes the Frobenius norm defined by the trace of $\boldsymbol{H}\boldsymbol{H}^t$. In the simulation, the $\{s_i\}$ are BPSK signals.

A. The Location of the CM Local Minimum

In the first experiment, SNR $= 10$ dB. From (24), the Wiener receivers are the columns of $\boldsymbol{\Phi}^{-1}$

$$\boldsymbol{\Phi}^{-1} = \begin{pmatrix} 0.5719 & 0.1527 \\ 0.1527 & 0.8882 \end{pmatrix} \Longrightarrow \boldsymbol{q}_m^{(1)} = \begin{pmatrix} 0.5719 \\ 0.1527 \end{pmatrix}$$

$$\boldsymbol{q}_m^{(2)} = \begin{pmatrix} 0.1527 \\ 0.8882 \end{pmatrix}. \tag{58}$$

The CMA receivers are obtained from the gradient search for the local minima of (20) initialized at the Wiener receivers.

The Contours of the Cost Function: Fig. 16 shows the contours of the CM cost function as well as the Wiener and CM receivers. At the origin, there is a local maximum of $J(\boldsymbol{q})$. The shape of $J(\boldsymbol{q})$ looks like a sombrero. The "hatband" (where the brim meets the crown) is the region containing the CM local minima. According to Theorem 1, the energy of a CM local minimum is between $1/3$ and 1 which are shown by two dashed lines in Fig. 16. This ring only roughly describes the location of CM minima and is not very tight. The signal subspace is the whole space R^2 in this example. From Fig. 16, it can be seen that the CM local minima are closer to the Wiener receivers than the zero-forcing receivers, and are approximately in the directions of the Wiener receivers.

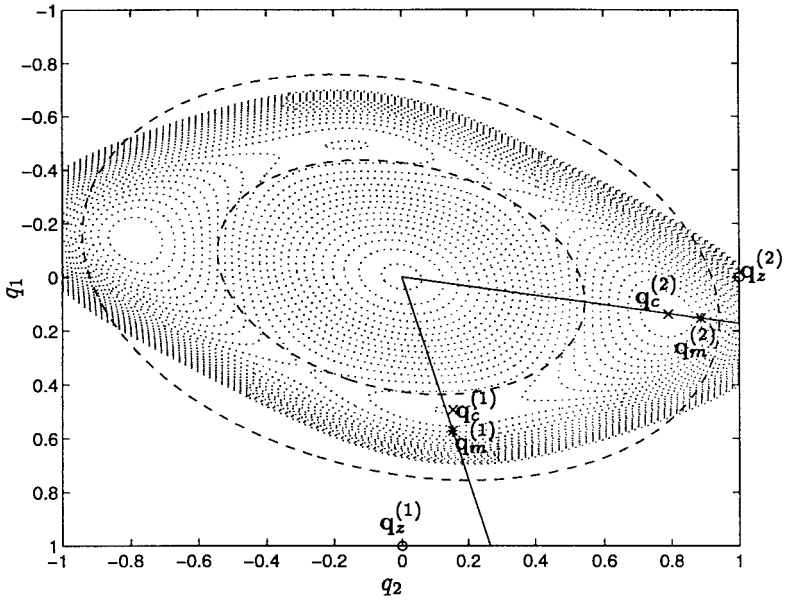

Fig. 16. The contours of the CM cost function. Cross: CM receiver; star: MMSE receiver; circle: ZF receiver.

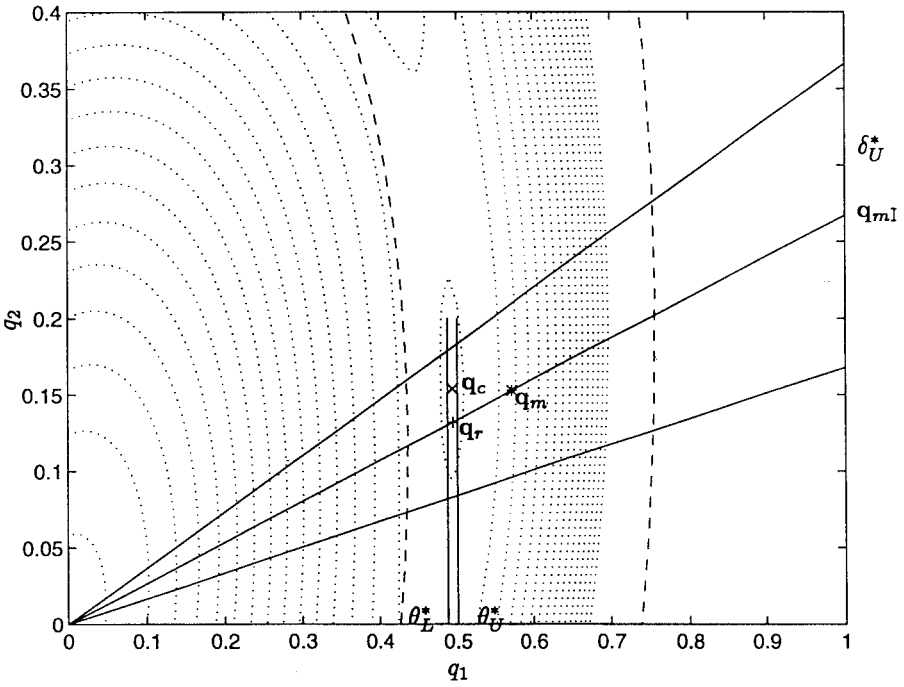

Fig. 17. The cone.

The Neighborhoods: First we determine the reference point $\boldsymbol{q}_r$ which has the minimum CM cost in the direction of $\boldsymbol{q}_m$. To obtain $\boldsymbol{q}_r$, one needs only the MSE and the ISI of the Wiener receiver $\boldsymbol{q}_m$. In this example, for $\nu = 1$, $\theta_m = 0.5719$, $\boldsymbol{q}_{m\mathrm{I}} = 0.1527/0.5719 = 0.2669$, and $||\boldsymbol{q}_{m\mathrm{I}}||_4 = 0.2669$. According to (37), we have

$$\theta_r = \sqrt{\frac{\theta_m}{3 - 2\theta_m^2 - 2\theta_m^2||\boldsymbol{q}_{m\mathrm{I}}||_4^4}} = 0.4941. \qquad (59)$$

The δ_U^*, θ_L^*, and θ_U^* are obtained according to Theorem 2. Specifically, δ_U^* is the smallest positive root of $D(\delta)$. Note that $J_m^\nu(\boldsymbol{q}_m^\nu) > 1/3$ for $\nu = 1$. But $c_2(\delta) > 0$ for $0 \le \delta \le \delta_U^*$. Theorem 2 is still applied to this case. Fig. 17 is a close-up of a segment of Fig. 16. The cone $\mathcal{K}_p$ in the two-dimensional space is a sector. The upper/lower bounds of θ_c turn out to be tight, while the upper bound of δ_c is loose.

In Table I, we compare the CM and the Wiener receivers in their MSE, output power, and ISI. Note that the MSE of CMA for estimating $s(k)$ $(\nu = 1)$ is about four times larger than that for estimating $s(k-1)$ $(\nu = 2)$. A similar ratio exists for the Wiener solutions with the two possible delays. Due to its blind initialization, CMA may perform considerably worse than a nonblind Wiener design with an optimal preselected delay.

B. The MSE Bound

In this experiment, we investigate the accuracy of the MSE upper bound $(\Delta\mathcal{E}_U)$ and the estimated MSE $\widehat{\Delta\mathcal{E}}$. We also

TABLE I
COMPARISONS OF THE CM AND THE WIENER RECEIVERS

	$\nu = 1$			$\nu = 2$		
	MSE	Power	ISI	MSE	Power	ISI
$\boldsymbol{q}_m^\nu$	0.4281	0.5719	0.2669	0.1118	0.8882	0.1719
$\boldsymbol{q}_c^\nu$	0.4393	0.4271	0.3113	0.1225	0.7039	0.1752

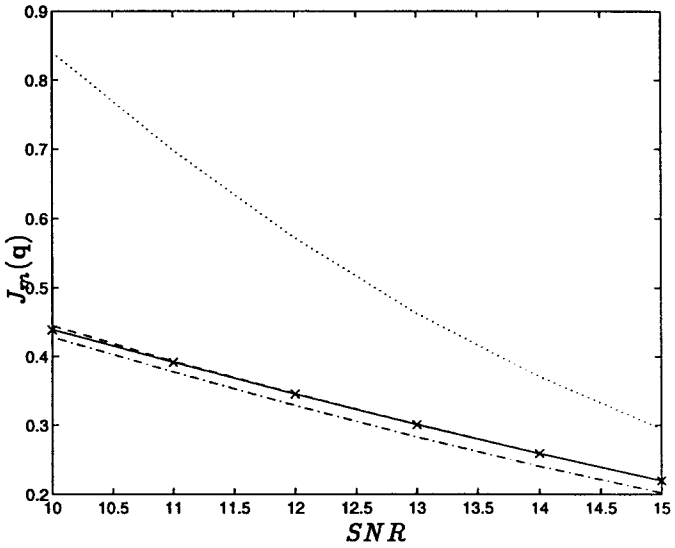

Fig. 18. Mean-squared errors. Solid line: $J_m(\boldsymbol{q}_c)$; dashed line: $J_m(\boldsymbol{q}_c) + \Delta\mathcal{E}_U$; dash–dot: $J_m(\boldsymbol{q}_m)$; cross: $J_m(\boldsymbol{q}_c) + \widehat{\Delta\mathcal{E}}$. Dotted line: $J_m(\boldsymbol{q}_c) + \widetilde{\Delta\mathcal{E}}$. $\nu = 1$.

compare the estimated MSE $\widehat{\Delta\mathcal{E}}$ to the value $\widetilde{\Delta\mathcal{E}}$ obtained from the second-order approximation [15], [20], i.e.,

$$\widetilde{\Delta\mathcal{E}} \triangleq \tfrac{9}{4} J_m^2(\boldsymbol{q}_m). \tag{60}$$

To examine the accuracy, we compare them with the actual MSE of CM receivers and the MMSE at different noise levels. The SNR varies from 10 to 15 dB.

Fig. 18 shows the comparison result. The upper bound is very tight. The estimated MSE $\widehat{\Delta\mathcal{E}}$, which lies between the upper bound $\Delta\mathcal{E}_U$ and the actual MSE $\mathcal{E}(\boldsymbol{q}_c)$, proves to be a good estimate and only requires two parameters for its computation (see (37) and (55)). However, the approximated MSE based on the second-order approximation [15], [20] has large error at low SNR.

VI. CONCLUSION

In this paper, we presented a geometrical analysis of the well-known CM receivers, which reveals connections between the (blind) CM and the (nonblind) Wiener receivers. In answering three questions posted in the Introduction, we obtain the following results:

C1: CM has local minima when noise is present. Interestingly, when the MMSE is less than $1/3$, the existence of a CM receiver in the neighborhood of the Wiener receiver can be established analytically by Theorem 2. CM local minima in other regions are not addressed in this paper. In other words, we have located "good" CM local minima with low MSE.

C2: The MSE of those "good" CM local minima is bounded in Theorem 4 using the magnitudes of the bias (B) and intersymbol/user interference (I) of the Wiener receiver. The upper bound (53) derived in this paper is tight in both the artificial and the practical examples (see [28]).

C3: We observe that the CM receiver is approximately the scaled version of the Wiener receiver. This also implies that the extra MSE of the CM receiver is simply given by

$$\left(\frac{1}{\sqrt{3-2(1-B)^2(1+I^4)}} - \sqrt{1-B}\right)^2. \tag{61}$$

Our analysis shows that the CM receivers can perform almost as well as the nonblind/trained receiver design if undesirable local minima can be avoided. How to achieve this by proper initialization remains one of the most important issue in the application of the CMA.

APPENDIX A
PROOF OF THEOREM 1

Given any $\boldsymbol{f} \in \mathcal{R}^{n_f}$, $\boldsymbol{f}$ can be decomposed by the projections in $\mathcal{R}_{\boldsymbol{H}}$ and $\mathcal{R}_{\boldsymbol{H}^\perp}$

$$\boldsymbol{f} = \boldsymbol{f}_{\boldsymbol{H}} + \boldsymbol{f}_{\boldsymbol{H}^\perp}. \tag{62}$$

Since

$$\boldsymbol{f}^t\boldsymbol{R}\boldsymbol{f} = \boldsymbol{f}_{\boldsymbol{H}}^t\boldsymbol{R}\boldsymbol{f}_{\boldsymbol{H}} + \boldsymbol{f}_{\boldsymbol{H}^\perp}^t\boldsymbol{R}\boldsymbol{f}_{\boldsymbol{H}^\perp} \tag{63}$$

the CM cost function (15) is

$$\begin{aligned}\overline{J}_c(\boldsymbol{f}) &= \overline{J}_c(\boldsymbol{f}_{\boldsymbol{H}}) + 3\|\boldsymbol{f}_{\boldsymbol{H}^\perp}\|_{\boldsymbol{R}}^4 \\ &\quad + 6\|\boldsymbol{f}_{\boldsymbol{H}}\|_{\boldsymbol{R}}^2\|\boldsymbol{f}_{\boldsymbol{H}^\perp}\|_{\boldsymbol{R}}^2 - 2\|\boldsymbol{f}_{\boldsymbol{H}^\perp}\|_{\boldsymbol{R}}^2 \\ &= \overline{J}_c(\boldsymbol{f}_{\boldsymbol{H}}) + 3\|\boldsymbol{f}_{\boldsymbol{H}^\perp}\|_{\boldsymbol{R}}^4 \\ &\quad + 6\|\boldsymbol{f}_{\boldsymbol{H}^\perp}\|_{\boldsymbol{R}}^2(\|\boldsymbol{f}_{\boldsymbol{H}}\|_{\boldsymbol{R}}^2 - \tfrac{1}{3}).\end{aligned} \tag{64}$$

Define $\mathcal{A} \triangleq \{\boldsymbol{f}\colon \|\boldsymbol{f}_{\boldsymbol{H}}\|_{\boldsymbol{R}}^2 > \frac{1}{3}\}$, $\forall\, \boldsymbol{f} \in \mathcal{A}$, $\overline{J}_c(\boldsymbol{f}) \geq \overline{J}_c(\boldsymbol{f}_{\boldsymbol{H}})$. Hence, the set of minimum points of $\overline{J}_c(\boldsymbol{f})(\forall \boldsymbol{f} \in \mathcal{A})$ is equivalent to $\overline{J}_c(\boldsymbol{f})(\forall \boldsymbol{f} \in \mathcal{R}_{\boldsymbol{H}} \bigcap \mathcal{A})$.

On the other hand, if $\boldsymbol{f} \in \mathcal{A}^c$ is a minimum point of $\overline{J}_c(\boldsymbol{f})$, then

$$\begin{aligned}\frac{\partial \overline{J}_c(\boldsymbol{f})}{\partial \boldsymbol{f}_{\boldsymbol{H}^\perp}} &= \frac{\partial \overline{J}_c(\boldsymbol{f})}{\partial \|\boldsymbol{f}_{\boldsymbol{H}^\perp}\|_{\boldsymbol{R}}^2}\frac{\partial \|\boldsymbol{f}_{\boldsymbol{H}^\perp}\|_{\boldsymbol{R}}^2}{\partial \boldsymbol{f}_{\boldsymbol{H}^\perp}} \\ &= (12\|\boldsymbol{f}_{\boldsymbol{H}^\perp}\|_{\boldsymbol{R}}^2 + 12(\|\boldsymbol{f}_{\boldsymbol{H}}\|_{\boldsymbol{R}}^2 - \tfrac{1}{3}))\boldsymbol{f}_{\boldsymbol{H}^\perp}^t\boldsymbol{R} \\ &= (12\|\boldsymbol{f}_{\boldsymbol{H}^\perp}\|_{\boldsymbol{R}}^2 + 12(\|\boldsymbol{f}_{\boldsymbol{H}}\|_{\boldsymbol{R}}^2 - \tfrac{1}{3}))\sigma^2\boldsymbol{f}_{\boldsymbol{H}^\perp}^t = 0.\end{aligned} \tag{65}$$

This implies that $\|\boldsymbol{f}\|_{\boldsymbol{R}}^2 = \frac{1}{3}$, or $\|\boldsymbol{f}_{\boldsymbol{H}^\perp}\| = 0$. If $\|\boldsymbol{f}\|_{\boldsymbol{R}}^2 = \frac{1}{3}$, from (15)

$$\overline{J}_c(\boldsymbol{f}) = \tfrac{2}{3} - 2\|\boldsymbol{H}^t\boldsymbol{f}\|_4^4. \tag{66}$$

If $\|\boldsymbol{f}_{\boldsymbol{H}^\perp}\| \neq 0$, one can find $\alpha > 1$, such that $\hat{\boldsymbol{f}} \triangleq \alpha\boldsymbol{f}_{\boldsymbol{H}}$ satisfies $\|\hat{\boldsymbol{f}}\|_{\boldsymbol{R}}^2 = 1/3$. Obviously, $\overline{J}_c(\boldsymbol{f}) > \overline{J}_c(\hat{\boldsymbol{f}})$. Thus $\overline{J}_c(\boldsymbol{f})$ is minimized when $\|\boldsymbol{f}_{\boldsymbol{H}^\perp}\| = 0$. Hence if $\boldsymbol{f} \in \mathcal{A}^c$ is a minimum point of (64), $\boldsymbol{f}$ must satisfy both conditions: $\|\boldsymbol{f}\|_{\boldsymbol{R}}^2 = \frac{1}{3}$ and $\|\boldsymbol{f}_{\boldsymbol{H}^\perp}\| = 0$. On the other hand, the gradient

of $\overline{J}_c(\boldsymbol{f})$ in the direction of this minimum point is given by

$$\frac{\partial \overline{J}_c(k\boldsymbol{f})}{\partial k}\Big|_{k=1} = 12||\boldsymbol{f}||_{\boldsymbol{R}}^4 - 4||\boldsymbol{f}||_{\boldsymbol{R}}^2 - 8||\boldsymbol{H}^t\boldsymbol{f}||_4^4 < 0. \quad (67)$$

This implies that there is no minimum point in $\mathcal{A}^c$. In other words, all CM local minimum points are in the signal subspace. At the same time, it also shows that the energy of minimum point is greater that $1/3$. To show the energy is less than 1, let us define a function for any $\boldsymbol{f}$ such that $||\boldsymbol{f}||_{\boldsymbol{R}} = 1$.

$$\phi(k, \boldsymbol{f}) \triangleq \overline{J}_c(\sqrt{k}\boldsymbol{f}) = (3 - 2||\boldsymbol{H}^t\boldsymbol{f}||_4^4)k^2 - 2k + 1. \quad (68)$$

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

$$k_{\min} = \frac{1}{3 - 2||\boldsymbol{H}^t\boldsymbol{f}||_4^4}. \quad (69)$$

Since

$$||\boldsymbol{H}^t\boldsymbol{f}||_4^4 \le ||\boldsymbol{H}^t\boldsymbol{f}||_2^4 < ||\boldsymbol{f}||_{\boldsymbol{R}}^4 = 1$$

$k_{\min} < 1$. This implies that the energy of the CMA equalizer cannot be greater than 1. □

Appendix B
Proof of Lemma 2

Since the noise variance $\sigma^2 > 0$, we have $\boldsymbol{C} > \boldsymbol{I}_{n_q-1}$, and

$$\begin{aligned} ||\boldsymbol{q}_{\mathrm{I}}||_4 &= ||\boldsymbol{q}_{\mathrm{I}} - \boldsymbol{q}_{m\mathrm{I}} + \boldsymbol{q}_{m\mathrm{I}}||_4 \le ||\boldsymbol{q}_{\mathrm{I}} - \boldsymbol{q}_{m\mathrm{I}}||_4 + ||\boldsymbol{q}_{m\mathrm{I}}||_4 \\ &< ||\boldsymbol{q}_{\mathrm{I}} - \boldsymbol{q}_{m\mathrm{I}}||_{\boldsymbol{C}} + ||\boldsymbol{q}_{m\mathrm{I}}||_4 = \delta + ||\boldsymbol{q}_{m\mathrm{I}}||_4. \end{aligned} \quad (70)$$

Using c) of Lemma 1 and the above inequality, we have

$$\begin{aligned} J(\boldsymbol{q}) - J(\boldsymbol{q}_r) &= 3(\boldsymbol{q}^t\boldsymbol{\Phi}\boldsymbol{q})^2 - 2(\boldsymbol{q}^t\boldsymbol{\Phi}\boldsymbol{q}) - 2||\boldsymbol{q}||_4^4 + 1 - J(\boldsymbol{q}_r) \\ &> 3\left(\delta^2 + \frac{1}{\theta_m}\right)^2\theta^4 - 2\left(\delta^2 + \frac{1}{\theta_m}\right)\theta^2 \\ &\quad - 2\theta^4(1 + (\delta + ||\boldsymbol{q}_{m\mathrm{I}}||_4)^4) + 1 - J(\boldsymbol{q}_r) \\ &= \underbrace{\left(3\left(\delta^2 + \frac{1}{\theta_m}\right)^2 - 2(1 + (\delta + ||\boldsymbol{q}_{m\mathrm{I}}||_4)^4)\right)}_{c_2(\delta)}\theta^4 \\ &\quad \underbrace{-2\left(\delta^2 + \frac{1}{\theta_m}\right)}_{c_1(\delta)}\theta^2 + \underbrace{1 - J(\boldsymbol{q}_r)}_{c_0}. \end{aligned} \quad (71)$$

Substituting (38) into $1 - J(\boldsymbol{q}_r)$, we obtain (45). □

Appendix C
Proof of Theorem 2

The outline of the proof is as follows. Consider the lower bound in (44). We first examine the sign of $c_2(\delta)$. From the signs of $c_2(\delta)$ and $D(\delta)$, one can determine the sign of polynomial (44). Second, we prove that all points on the peripheral surface $\mathcal{S}_p$ (see Fig. 19) have costs greater than the reference point. The points on the upper surface $\mathcal{S}_u$ and the lower surface $\mathcal{S}_L$ will be checked next.

Finally, we verify that the reference point is in the sliced cone.

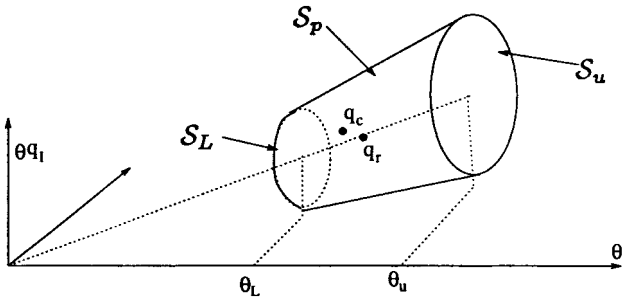

Fig. 19. Three surfaces of the sliced cone.

1) We prove that $c_2(\delta) > 0$. Since that $0 \le \delta < ||\boldsymbol{q}_{m\mathrm{I}}||_2$, we have

$$\begin{aligned} c_2(\delta) &= 3\left(\delta^2 + \frac{1}{\theta_m}\right)^2 - 2(1 + (\delta + ||\boldsymbol{q}_{m\mathrm{I}}||_4)^4) \quad (72) \\ &> 3\left(\delta^2 + \frac{1}{\theta_m}\right)^2 \\ &\quad - 2(1 + \delta^4 + 6||\boldsymbol{q}_{m\mathrm{I}}||_2^2\delta^2 + 9||\boldsymbol{q}_{m\mathrm{I}}||_2^4) \quad (73) \\ &> \delta^4 + \left(\frac{6}{\theta_m} - 12||\boldsymbol{q}_{m\mathrm{I}}||_2^2\right)\delta^2 \\ &\quad + \left(\frac{3}{\theta_m^2} - 2 - 18||\boldsymbol{q}_{m\mathrm{I}}||_2^4\right) \quad (74) \\ &> \delta^4 + \left(\frac{6}{\theta_m} - 12\frac{1 - \theta_m}{\theta_m}\right)\delta^2 \\ &\quad + \left(\frac{3}{\theta_m^2} - 2 - 18\left(\frac{1 - \theta_m}{\theta_m}\right)^2\right). \quad (75) \end{aligned}$$

For (75), we have used the property b) in Lemma 1.

$$||\boldsymbol{q}_{m\mathrm{I}}||_2^2 = \frac{||\theta_m\boldsymbol{q}_{m\mathrm{I}}||^2}{\theta_m^2} < \frac{J_m(\boldsymbol{q}_m) - (1 - \theta_m)^2}{\theta_m^2} = \frac{1 - \theta_m}{\theta_m}. \quad (76)$$

When $J_m(\boldsymbol{q}_m) < 1/3$ or $\theta_m \in (\frac{2}{3}, 1)$ the last two terms of (75) are positive.

2) Now we evaluate the cost on the peripheral surface $\mathcal{S}_p$ defined by $\partial\mathcal{K}_p(\boldsymbol{q}_m, \delta_U^*)$. From

$$\begin{aligned} D(\delta) &= 4\left(\delta^2 + \frac{1}{\theta_m}\right)^2 \\ &\quad - 4\frac{3\left(\delta^2 + \frac{1}{\theta_m}\right)^2 - 2(1 + (\delta + ||\boldsymbol{q}_{m\mathrm{I}}||_4)^4)}{3 - 2\theta_m^2 - 2\theta_m^2||\boldsymbol{q}_{m\mathrm{I}}||_4^4} \end{aligned} \quad (77)$$

one can easily see that $D(0) = 0$ and $dD(0)/d\delta > 0$. If $D(||\boldsymbol{q}_{m\mathrm{I}}||_2) < 0$, then there exists $\delta_U^* \in (0, ||\boldsymbol{q}_{m\mathrm{I}}||_2)$ such that $D(\delta_U^*) \le 0$. Since $c_2(\delta_U^*) > 0$, the polynomial

$$c_2(\delta_U^*)\theta^4 + c_1(\delta_U^*)\theta^2 + c_0 \ge 0$$

for all θ^2. Thus $\forall\, \boldsymbol{q} \in \partial\mathcal{K}_p(\boldsymbol{q}_m, \delta_U^*)$, $J(\boldsymbol{q}) - J(\boldsymbol{q}_r) > 0$.

3) Now we check the points on the upper surface $\mathcal{S}_u$ defined by $\theta = \theta_U^*, 0 \le \delta \le \delta_U^*$. For all points on this surface, $c_2(\delta) > 0$. Since

$$\theta^2 = (\theta_U^*)^2 \ge \frac{-c_1(\delta) + \sqrt{c_1(\delta)^2 - 4c_2(\delta)c_0}}{2c_2(\delta)} \quad (78)$$

then the polynomial $c_2(\delta)\theta^4 + c_1(\delta)\theta^2 + c_0 \geq 0$. Hence, $J(\boldsymbol{q}) - J(\boldsymbol{q}_r) > 0, \forall 0 < \delta \leq \delta_U^*$. In other words, we prove that all points except the point on the line $\boldsymbol{q}_m$ have costs greater than $J(\boldsymbol{q}_r)$. Similarly, we can prove the same result for the lower surface.

4) Finally, we verify that the reference point $\boldsymbol{q}_r$ is in the region $\mathcal{B}_p(\boldsymbol{q}_m, \delta_U^*, \theta_L^*, \theta_U^*)$, i.e., $\theta_L^* < \theta_r < \theta_U^*$. Since $\boldsymbol{q}_r$ is the minimum point on the line $\delta = 0$, consequently, the costs at $\delta = 0$, $\theta = \theta_U^*$, and $\theta = \theta_L^*$ are greater than $J(\boldsymbol{q}_r)$. It is easy to show that

$$\frac{-c_1(\delta) + \sqrt{c_1(\delta)^2 - 4c_2(\delta)c_0}}{2c_2(\delta)} \tag{79}$$

is increasing at $\delta = 0$. From (50), we have

$$(\theta_U^*)^2 > \frac{-c_1(0) + \sqrt{c_1(0)^2 - 4c_2(0)c_0}}{2c_2(0)} \tag{80}$$

$$\geq \frac{-c_1(0)}{2c_2(0)} = \frac{\theta_m}{3 - 2\theta_m^2 - 2\theta_m^2 ||\boldsymbol{q}_{m\mathrm{I}}||_4^4} = \theta_r^2. \tag{81}$$

Similarly, from (49), we have

$$(\theta_L^*)^2 < \frac{-c_1(0) - \sqrt{c_1(0)^2 - 4c_2(0)c_0}}{2c_2(0)} = \frac{-c_1(0)}{2c_2(0)} = \theta_r^2. \tag{82}$$

In the end, we examine the case as the noise variance $\sigma \to 0$. From (45)–(47), we have

$$\lim_{\sigma \to 0} c_0 = 1 \tag{83}$$

$$\lim_{\sigma \to 0} c_1(\delta) = -2(\delta^2 + 1) \tag{84}$$

$$\lim_{\sigma \to 0} c_2(\delta) = 3(\delta^2 + 1)^2 - 2(1 + \delta^4). \tag{85}$$

Hence, $D(\delta) = -16\delta^2 < 0$. Therefore, $\delta_U^* = 0$, and $\theta_L^* = \theta_U^* = 1 = \theta_m$. □

Appendix D
Proof of Theorem 3

a) In order to prove this theorem, we show that there does not exist a CM local minimum in the cone $\mathcal{K}_p(\boldsymbol{q}_m, ||\boldsymbol{q}_{m\mathrm{I}}||_2)$ with power greater than that of the Wiener receiver.

Any point in $\mathcal{K}_p(\boldsymbol{q}_m, ||\boldsymbol{q}_{m\mathrm{I}}||_2)$ can be represented by $\sqrt{k}\boldsymbol{q}$, where $||\boldsymbol{q}||_{\boldsymbol{\Phi}}^2 = ||\boldsymbol{q}_m||_{\boldsymbol{\Phi}}^2 = \theta_m$, and $0 \leq k < \infty$. Define

$$\phi(k, \boldsymbol{q}) \triangleq \tfrac{1}{2} J(\sqrt{k}\boldsymbol{q}) = \tfrac{1}{2}(3\theta_m^2 k^2 - 2\theta_m k - 2||\boldsymbol{q}||_4^4 k^2 + 1). \tag{86}$$

Since $\theta_m = ||\boldsymbol{q}||_{\boldsymbol{\Phi}}^2 > ||\boldsymbol{q}||_2^2 \geq ||\boldsymbol{q}||_4^2$, the derivative

$$\phi'(k, \boldsymbol{q}) \triangleq \frac{d}{dk}\phi(k, \boldsymbol{q}) = (3\theta_m^2 - 2||\boldsymbol{q}||_4^4)k - \theta_m \tag{87}$$

is monotonically increasing with $\phi'(0, \boldsymbol{q}) < 0$. Therefore, if $\phi'(1, \boldsymbol{q}) > 0$, then there is no CM receiver of the form $k\boldsymbol{q}$ for any $k > 1$.

Now we prove that $\forall\, \boldsymbol{q} \in \mathcal{K}_p(\boldsymbol{q}_m, ||\boldsymbol{q}_{m\mathrm{I}}||_2) \bigcap \{||\boldsymbol{q}||_{\boldsymbol{\Phi}}^2 = \theta_m\}$, $\phi'(1, \boldsymbol{q}) > 0$. Since

$$||\boldsymbol{q}||_4^4 = \theta^4 + ||\theta \boldsymbol{q}_{\mathrm{I}}||_4^4 \leq \theta^4 + ||\theta \boldsymbol{q}_{\mathrm{I}}||_2^4$$
$$\leq \theta^4 + (\boldsymbol{q}^t \boldsymbol{\Phi} \boldsymbol{q} - \theta^2)^2 = \theta^4 + (\theta_m - \theta^2)^2 \tag{88}$$

we have

$$\phi'(1, \boldsymbol{q}) = (3\theta_m - 1)\theta_m - 2||\boldsymbol{q}||_4^4$$
$$\geq (3\theta_m - 1)\theta_m - 2\theta^4 - 2(\theta_m - \theta^2)^2$$
$$= -4\theta^4 + 4\theta_m \theta^2 + \theta_m^2 - \theta_m. \tag{89}$$

If

$$\theta_m\left(\frac{1}{2} - \frac{1}{2}\sqrt{2 - \frac{1}{\theta_m}}\right) < \theta^2 < \theta_m\left(\frac{1}{2} + \frac{1}{2}\sqrt{2 - \frac{1}{\theta_m}}\right) \tag{90}$$

then $\phi'(1, \boldsymbol{q}) > 0$.

Now we check whether the condition (90) is satisfied for all points in

$$\mathcal{K}_p(\boldsymbol{q}_m, ||\boldsymbol{q}_{m\mathrm{I}}||_2) \bigcap \{||\boldsymbol{q}||_{\boldsymbol{\Phi}}^2 = \theta_m\}.$$

According to c) of Lemma 1, $\boldsymbol{q}^t \boldsymbol{\Phi} \boldsymbol{q} = \theta^2(\delta^2 + (1/\theta_m)) = \theta_m$, hence

$$\theta^2 = \frac{\theta_m}{\delta^2 + \dfrac{1}{\theta_m}}. \tag{91}$$

Using b) of Lemma 1, we have

$$||\boldsymbol{q}_{m\mathrm{I}}||_2^2 = \frac{||\theta_m \boldsymbol{q}_{m\mathrm{I}}||^2}{\theta_m^2} < \frac{J_m(\boldsymbol{q}_m) - (1 - \theta_m)^2}{\theta_m^2} = \frac{1 - \theta_m}{\theta_m}. \tag{92}$$

Using d) of Lemma 1, we thus have

$$\frac{\theta_m^2}{2 - \theta_m} \leq \frac{\theta_m}{||\boldsymbol{q}_{m\mathrm{I}}||_2^2 + \dfrac{1}{\theta_m}} \leq \theta^2 \leq \theta_m^2. \tag{93}$$

After straightforward manipulations, one can show that, when $\theta_m > 0.6$

$$\theta_m\left(\frac{1}{2} - \frac{1}{2}\sqrt{2 - \frac{1}{\theta_m}}\right) < \frac{\theta_m^2}{2 - \theta_m} \tag{94}$$

$$\theta_m^2 < \theta_m\left(\frac{1}{2} + \frac{1}{2}\sqrt{2 - \frac{1}{\theta_m}}\right). \tag{95}$$

This completes the proof that the output power of the CM receiver is less than $||\boldsymbol{q}_m||_{\boldsymbol{\Phi}}^2$. From d) of Lemma 1, we have $\theta_c \leq \theta_m$, i.e., the bias of the CM receiver is greater that of Wiener receiver.

b) Prove by contradiction. Suppose that $||\boldsymbol{q}_{c\mathrm{I}}||_4 < ||\boldsymbol{q}_{m\mathrm{I}}||_4$. Let

$$\hat{\boldsymbol{q}}_c \triangleq \theta_c \begin{pmatrix} 1 \\ \boldsymbol{q}_{m\mathrm{I}} \end{pmatrix}. \tag{96}$$

Obviously, $\hat{\boldsymbol{q}}_c \in \mathcal{B}_p(\boldsymbol{q}_m, \delta_U^*, \theta_L^*, \theta_U^*)$, and $||\hat{\boldsymbol{q}}_c||_4 > ||\boldsymbol{q}_c||_4$. According to (29)

$$||\hat{\boldsymbol{q}}_c||_{\boldsymbol{\Phi}}^2 = \frac{\theta_c^2}{\theta_m} \leq \left(\delta_c^2 + \frac{1}{\theta_m}\right)\theta_c^2 \leq ||\boldsymbol{q}_c||_{\boldsymbol{\Phi}}^2. \tag{97}$$

Therefore,

$$J(\hat{\boldsymbol{q}}_c) < J(\boldsymbol{q}_c) \tag{98}$$

which contradicts that $\boldsymbol{q}_c$ is the CM minimum in $\mathcal{B}_p(\boldsymbol{q}_m, \delta_U^*, \theta_L^*, \theta_U^*)$. □

APPENDIX E
PROOF OF THEOREM 4

From (20) and c) of Lemma 1

$$\begin{aligned}\Delta\mathcal{E} &= J_m\left(\theta_c\begin{pmatrix}1\\ \boldsymbol{q}_{c\mathrm{I}}\end{pmatrix}\right) - J_m(\boldsymbol{q}_m)\\ &= 1 - 2\theta_c + \theta_c^2\left(\delta_c^2 + \frac{1}{\theta_m}\right) - (1-\theta_m)\\ &= \frac{(\theta_c-\theta_m)^2}{\theta_m} + \theta_c^2\delta_c^2. \end{aligned} \tag{99}$$

Thus we obtain the bounds in the theorem. This relationship is also evident in Fig. 15 where

$$\Delta\mathcal{E} = E\{(y_c - y_m)^2\} = E\{(y_c - x)^2\} + E\{(y_m - x)^2\}.$$

To show (55), we approximate the CM receiver $\boldsymbol{q}_c$ by the reference $\boldsymbol{q}_r = \theta_r\binom{1}{\boldsymbol{q}_{m\mathrm{I}}}$. Hence, using c) of Lemma 1 or Fig. 15

$$\widehat{\Delta\mathcal{E}} = E\{(y_r - y_m)^2\} = \frac{(\theta_r-\theta_m)^2}{\theta_m} = \left(\frac{\theta_r}{\sqrt{\theta_m}} - \sqrt{\theta_m}\right)^2. \tag{100}$$

According to (37)

$$\begin{aligned}\frac{\theta_r}{\sqrt{\theta_m}} &= (1 + 2(1-\theta_m^2(1+||\boldsymbol{q}_{m\mathrm{I}}||_4^4)))^{-(1/2)}\\ &= 1 - (1-\theta_m^2(1+||\boldsymbol{q}_{m\mathrm{I}}||_4^4))\\ &\quad + O(1-\theta_m^2(1+||\boldsymbol{q}_{m\mathrm{I}}||_4^4)). \end{aligned} \tag{101}$$

Since

$$\begin{aligned}1 - \theta_m^2(1+||\boldsymbol{q}_{m\mathrm{I}}||_4^4) &= (1+\theta_m)(1-\theta_m) - \theta_m^2||\boldsymbol{q}_{m\mathrm{I}}||_4^4\\ &= 2J_m(\boldsymbol{q}_m) + O(J_m(\boldsymbol{q}_m)) \end{aligned} \tag{102}$$

thus

$$\begin{aligned}\frac{\theta_r}{\sqrt{\theta_m}} - \sqrt{\theta_m} &= [1 - 2J_m(\boldsymbol{q}_m)] - [1 - \tfrac{1}{2}J_m(\boldsymbol{q}_m)]\\ &\quad + O(J_m(\boldsymbol{q}_m)). \end{aligned} \tag{103}$$

Therefore,

$$\widehat{\Delta\mathcal{E}} = \tfrac{9}{4}J_m^2(\boldsymbol{q}_m) + O(J_m^2(\boldsymbol{q}_m)). \tag{104}$$

□

ACKNOWLEDGMENT

The authors gratefully acknowledge the comments from Phil Schniter, Raúl Casas, and Fernando López de Victoria.

REFERENCES

[1] J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu, and G. D. Forney, Jr., "MMSE decision feedback equalization and coding—Part I and II," *IEEE Trans. Commun.*, vol. 43, pp. 2582–2604, Oct. 1995.

[2] Z. Ding and C. R. Johnson, Jr.,"On the nonvanishing stability of undesirable equilibria for FIR Godard blind equalizers," *IEEE Trans. Signal Processing*, vol. 42, pp. 1940–1944, May 1993.

[3] 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. 41, pp. 2425–2432, Oct. 1992.

[4] Z. Ding, R. A. Kennedy, B. 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.

[5] ——, "Local convergence of the Sato blind equalizer and generalizations under practical constraints," *IEEE Trans. Inform. Theory*, vol. 39, pp. 128–144, Jan. 1993.

[6] T. Ericson, "Structure of optimum receiving filters in data transmission systems," *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 352–353, May 1971.

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

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

[9] R. P. Gooch and J. D. Lundell, "The CM array: An adaptive beamformer for constant modulus signals," in *Proc. ICASSP'86 Conf.* (Tokyo, Japan, Apr. 1986), pp. 2523–2526.

[10] H. Zeng and L. Tong, "Mean-squared error performance of constant modulus receiver for singular channels," in *Proc. 1997 IEEE Int. Conf. Acoustics, Speech, and Signal Processing* (Munich, Germany, Apr. 1997), vol. I, pp. 647–650.

[11] C. R. Johnson, Jr., "Yet still more on the interaction of adaptive filtering, identification and control," *IEEE Signal Processing Mag.*, vol. 12, pp. 22–37, Mar. 1995.

[12] 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.

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

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

[15] Y. Li, K. Liu, and Z. Ding, "Length-dependent and cost-dependent local minima of unconstrained blind channel equalizers," *IEEE Trans. Signal Processing*, vol. 44, p. 2726, Nov. 1996.

[16] D. Liu and L. Tong, "An analysis of constant modulus algorithm for array signal processing," *Signal Processing*, to be published.

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

[18] J. Proakis, *Digital Communications.* New York: McGraw-Hill, 1995.

[19] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing*, 2nd ed. London, U.K.: Macmillan, 1992.

[20] H. Zeng, S. Zeng, and L. Tong, "Blind equalization using CMA: Performance analysis and a new algorithm," in *Proc. 1996 IEEE Int. Conf. Communications*, 1996.

[21] L. Tong, G. Xu, B. Hassibi, and T. Kailath, "Blind identification and equalization of multipath channels: A frequency domain approach," *IEEE Trans. Inform. Theory*, vol. 41, pp. 329–334, Jan. 1995.

[22] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inform. Theory*, vol. 40, pp. 340–349, Mar. 1994.

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

[24] 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.

[25] 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. 45–81, May 1996.

[26] P. P. Vaidyanathan, *Multirate Systems and Filter Banks.* Englewood Cliffs, NJ: Prentice Hall, 1993.

[27] H. Zeng, "Blind estimation: Performance analysis and new algorithms," Ph.D. dissertation, Univ. of Connecticut, Storrs, May 1997.

[28] 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. Information Sciences and Systems* (Baltimore, MD, Mar. 1996).

[29] H. Zeng, L. Tong, and C. R. Johnson, Jr., "An analysis of constant modulus receivers," *IEEE Trans. Signal Processing*, submitted for publication.

[30] ——, "Behavior of fractionally-spaced constant modulus algorithm, mean square error, robustness and local minima," in *Proc. 30th Asilomar Conf. on Signals, Systems, and Computers* (Pacific Grove, CA, Nov. 1996), vol. II, pp. 305–309.

[31] C. R. Johnson, Jr. *et al.*, "Blind equalization using the constant modulus criterion: A review," *Proc. IEEE*, vol. 86, Nov. 1998, to be published.