Geometrical Characterizations of Constant Modulus Receivers

Ming Gu, *Student Member, IEEE*, and Lang Tong, *Member, IEEE*

***Abstract*— Convergence properties of the constant modulus (CM) and the Shalvi–Weinstein (SW) algorithms in the presence of noise remain largely unknown. A new geometrical approach to the analysis of constant modulus and Shalvi–Weinstein receivers is proposed by considering a special constrained optimization involving norms of the combined channel-receiver response. This approach provides a unified framework within which various blind and (nonblind) Wiener receivers can all be analyzed by circumscribing an ellipsoid by norm balls of different types. A necessary and sufficient condition for the equivalence among constant modulus, Shalvi–Weinstein, zero forcing, and Wiener receivers is obtained. Answers to open questions with regard to CM and SW receivers, including their locations and their relationship with Wiener receivers, are provided for the special orthogonal channel and the general two-dimensional (2-D) channel-receiver impulse response. It is also shown that in two dimensions, each CM or SW receiver is associated with one and only one Wiener receiver.**

***Index Terms*— Adaptive filters, blind equalization, deconvolution.**

I. Introduction

A. The Problem

CONSIDER the linear estimation problem shown in Fig. 1:

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

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

where

- $\mathbf{s}$ source vector;
- $\mathbf{H}$ channel impulse response matrix;
- $\mathbf{w}$ additive white Gaussian noise;
- $\mathbf{x}$ received signal;
- y estimate of s_i (the ith element of $\mathbf{s}$) by a linear estimator $\mathbf{f}$.

The combined channel-receiver response is denoted by $\mathbf{q} \triangleq \mathbf{H}^t\mathbf{f}$. This model has a wide range of important applications in channel equalization, sensor array processing, multiuser detection, and source separation.

In designing the receiver coefficients, the well-known Wiener (MMSE) receiver is defined by minimizing the mean square error (MSE)

$$\overline{J}_m{}^{(i)}(\mathbf{f}) \triangleq E\{|y - s_i|^2\}. \tag{3}$$

This, however, requires the *joint* second-order moment of $\mathbf{x}$ and $\mathbf{s}$. When the channel input is inaccessible and the channel matrix is unknown, the receiver must be designed "blindly," using only marginal statistics of $\mathbf{s}$ and $\mathbf{x}$. One of the most popular blind approaches is the constant modulus algorithm (CMA) [6], [16], which minimizes the dispersion of the receiver output around some constant r. Specifically, a constant modulus (CM) receiver is a local minimum of the CM cost

$$\overline{J}_c(\mathbf{f}) \triangleq E\{(|y|^2 - r)^2\}. \tag{4}$$

Another criterion developed by Shalvi and Weinstein in [14] involves maximizing

$$\overline{J}_s(\mathbf{f}) \triangleq |K_4(y)| \tag{5}$$

subject to the power constraint $E\{|y|^2\} = E\{|s|^2\}$, where $K_4(y)$ is the fourth-order cumulant of y. Local maxima of the above constrained optimization are referred to as Shalvi–Weinstein (SW) receivers. In order to achieve global convergence, Vembu *et al.* [19] proposed a class of convex functions for blind receiver designs. Referred to as a VVK receiver, it is the minimum of the cost function

$$\overline{J}_v(\mathbf{f}) \triangleq \frac{1}{M} \lim_{p\to\infty} (E\{|y|^p\})^{1/p} \tag{6}$$

under the constraint $\mathbf{e}_i^t\mathbf{f} = 1$, where $M = \max\{|s_i|\}$ and $\mathbf{e}_i$ is the ith column of an identity matrix.

The constant modulus algorithm has been applied successfully to channel equalization and array signal processing. See recent surveys [3], [17], [18]. In his original paper [6], Godard observed that CM receivers appear to have a mean square error close to that of Wiener receivers. Similar observations have been widely reported leading to a formal analysis in [20]. For the Shalvi–Weinstein algorithm, it has been shown in [10] that optima of the SW cost are colinear with, and one-one correspondent to, those of the CM cost function, which implies that SW and CM receivers have the same asymptotic performance. Despite the increasing evidence, both in theory and in applications, that CM and SW receivers have Wiener-receiver-like performance, the following fundamental problems about their convergence properties remain unsolved.

- Q1) How many local optima are there?
- Q2) Where are these optimal solutions?
- Q3) Is every CM or SW receiver associated with a corresponding Wiener receiver?

Manuscript received May 13, 1998; revised April 8, 1999. This work was supported in part by the National Science Foundation under Contract NCR-9321813, the Office of Naval Research under Contract N00014-96-1-0895, and the Advanced Research Projects Agency, monitored by the Federal Bureau of Investigation, under Contract J-FBI-94-221. The associate editor coordinating the review of this paper and approving it for publication was Dr. Frans M. Coetzee.

M. Gu is with the Department of Electrical and Systems Engineering, University of Connecticut, Storrs, CT 06269 USA.

L. Tong is with the School of Electrical Engineering, Cornell University, Ithaca, NY 14853 USA.

Publisher Item Identifier S 1053-587X(99)07659-X.

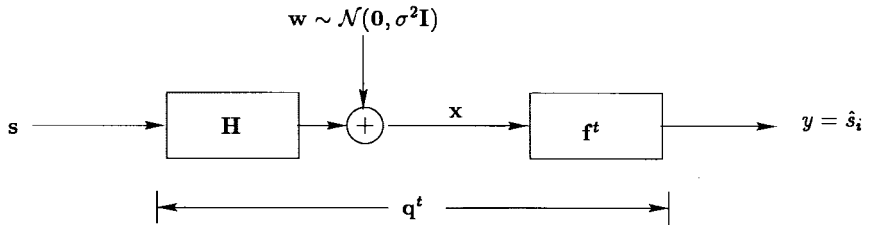

Fig. 1. Linear estimation model.

To gain insight into the difficulty of these questions, consider the direct approach to the analysis of CM receivers for the two-dimensional (2-D) case. The CM cost function (under assumptions listed in Section II) is shown in [20] to have the form

$$\overline{J}_c(\mathbf{f}) = 3\|\mathbf{f}\|_{\mathbf{R}}^4 - 2\|\mathbf{f}\|_{\mathbf{R}}^2 - 2\|\mathbf{f}^t\mathbf{H}\|_4^4 + 1 \tag{7}$$

where $\mathbf{R}$ is the covariance matrix of $\mathbf{x}$. Minimizing $\overline{J}_c(\mathbf{f})$ implies finding its stationary points by solving $\nabla\overline{J}_c(\mathbf{f}) = \mathbf{0}$, i.e.,

$$\nabla\overline{J}_c(\mathbf{f}) = 12\mathbf{f}^t\mathbf{R}\mathbf{f}\mathbf{R}\mathbf{f} - 4\mathbf{R}\mathbf{f} - 8\sum_{i=1}^{n}(\mathbf{f}^t\mathbf{H}\mathbf{e}_i)^3\mathbf{H}\mathbf{e}_i = \mathbf{0}. \tag{8}$$

In the 2-D case, (8) leads to two coupled third-order polynomial equations

$$\begin{cases} a_1f_1^3 + a_2f_1^2f_2 + a_3f_1f_2^2 + a_4f_2^3 + a_5f_1 + a_6f_2 = 0 \\ b_1f_1^3 + b_2f_1^2f_2 + b_3f_1f_2^2 + b_4f_2^3 + b_5f_1 + b_6f_2 = 0 \end{cases}$$

where a_i and b_i $(i = 1, \cdots, 6)$ can be calculated from $\mathbf{R}$ and $\mathbf{H}$. It is clear that answering Q1)–Q3) through the above algebraic equations is difficult if not impossible. In general, there is no closed-form solution.

B. Contributions and Related Work

In this paper, we present results in three areas. First, a new geometrical approach is developed to analyze several existing blind equalizers including CM, SW, and VVK receivers. Conventional zero forcing (ZF) and Wiener receivers can also be analyzed within the same framework. More importantly, connections among Wiener and blind receivers become apparent. The key of this approach is to convert various criteria to the problem of circumscribing an ellipsoid (l_2-norm ball) by different norm balls. Second, we show in Theorem 1 that CM, SW, ZF, and Wiener receivers are colinear and one-one correspondent *if and only if* either there is no noise or the channel is orthogonal. This result answers Q1)–Q3) for the orthogonal channels that are particularly important in the CDMA and OFDM applications [7], [8]. Third, we consider the general 2-D channel-receiver impulse response and provide solutions to the three open problems. Given the signal-to-interference ratio (SIR) of Wiener receivers (without knowing the channel), a necessary and sufficient condition is derived to determine the number of local CM and SW optima. Their locations are also obtained as functions of the SIR's of Wiener receivers.

Whereas this paper focuses on the global convergence behavior of CM and SW receivers, recent results [12], [20] on the local convergence property of CM receivers motivated our work. In [20], it was shown that when the MSE of a Wiener receiver is sufficiently small, there exists a CM receiver in the vicinity of the Wiener receiver. This local result is valid only around those Wiener receivers with reasonably small MSE. What has not been answered is the following: Is every CM receiver also associated with a corresponding Wiener receiver? Specifically, when the condition of the low MSE is not satisfied, are there any other CM receivers? If the answer is "Yes," are they still adjacent to the Wiener receivers or located somewhere else? The converse of the main theorem in [20] is established in this paper for the orthogonal channel and the general 2-D cases: There exists one and only one Wiener receiver in the neighborhood of every CM/SW[1] receiver. Furthermore, this global result provides the necessary and sufficient conditions to test whether and when a CM/SW receiver will appear near the Wiener receiver and tells how close they would be. The authors of [12] have also observed in simulation that there may be no CM receiver near a Wiener receiver. In Section IV, we prove, for the 2-D case, that if the CM/SW receiver for s_i does not exist in a certain neighborhood of the Wiener receiver for the same signal, there is no CM/SW receiver for s_i. This implies that CM and SW cost functions do not create spurious local optima not related to the Wiener receivers. Moreover, the existence and the locations of CM/SW receivers are determined by Wiener receivers. Another new property of CM/SW receivers in two dimensions is that at low signal-to-noise ratio (SNR), as the channel becomes more ill-conditioned, the performance of the global CM/SW receiver is closer to the corresponding Wiener receiver. This is somewhat surprising as one may expect that the loss of diversity at low SNR may have adverse effects on all CM receivers. All of the above results can be clearly explained by using the proposed geometrical approach, although, in this paper, it is intended for the analysis of the sub-Gaussian sources only. The role of the statistics of sources was studied, and the local results in [20] were generalized to the complex CMA in [12]. While extending the results in this paper to the general high-dimensional (including complex) case is not straightforward, this approach has been successfully applied to the independent component analysis of heterogeneous sources [15].

Attempts have been made over the years to answer Q1)–Q3). For the noiseless case, answers were given in [5] and [14]. Specifically, when $\mathbf{H}$ has full column rank, CM, SW, and Wiener receivers are identical. The error surface of the CM

[1] In this paper, wherever "CM/SW" is used, the conditions or the results stated apply to both CM and SW receivers.

cost function was analyzed in [9]. In the presence of noise, limited global results have been obtained. Of particular relevance to our work is the one-one correspondence between SW and CM receivers, proved by Li and Ding in [10]. Since SW receivers can be characterized using l_2-norm and l_4-norm balls, this correspondence enables us to analyze CM receivers geometrically. More recently, Chung and LeBlanc investigated the effects of noise on the number and the locations of local minima of the CM cost function [2]. It was shown that the number of CM local minima may be reduced as the noise power increases. In general, answers to Q1)–Q3) have not been reported even for the 2-D case. Local results have been obtained by two approaches. An *exact* analysis was proposed by Zeng *et al.* [20] that distinguishes those CM receivers with low MSE. On the other hand, by perturbing the CM cost function around either ZF or Wiener receivers, *approximate* characterizations of CM receivers can be obtained [4], [11], [21].

C. Organization and Notation

The rest of this paper is organized as follows. A new geometrical approach is proposed in Section II for the analysis of CM and SW receivers. In Section III, a necessary and sufficient condition on the equivalence among CM, SW, ZF, and Wiener receivers is presented. The locations of CM and SW receivers along with their close relationship with Wiener receivers are shown for the general 2-D case in Section IV. Several new properties of CM/SW receivers are observed, and numerical examples are given as well. The conclusion is drawn in Section V.

The notations used in this paper are standard. Uppercase and lowercase bold letters denote matrices and vectors, respectively. Key symbols are described in the following list.

- $(\cdot)^t$ Transpose.
- $(\cdot)^\dagger$ Pseudoinverse.
- $E\{\cdot\}$ Expectation operator.
- $\|\mathbf{x}\|_p$ l_p-norm of $\mathbf{x}$ defined by $(\sum_i |x_i|^p)^{1/p}$.
- $\|\mathbf{x}\|_\infty$ l_∞-norm of $\mathbf{x}$ defined by $\max_i |x_i|$.
- $\|\mathbf{x}\|_\mathbf{A}$ Vector norm of $\mathbf{x}$ on matrix $\mathbf{A}$ defined by $\sqrt{\mathbf{x}^t\mathbf{A}\mathbf{x}}$.
- $\mathbf{I}$ Identity matrix.
- $\mathbf{e}_i$ Unit column vector with 1 at the ith entry and zero elsewhere.
- $\chi(\mathbf{A})$ Condition number of matrix $\mathbf{A}$.
- $\angle\mathbf{x}$ Angle of 2-D vector $\mathbf{x} = \binom{x_1}{x_2}$ defined by $\arctan(x_2/x_1)$.
- ∇f Gradient of function f.

II. A Geometrical Approach

We make the following assumptions in our analysis.

- A1) Entries of $\mathbf{s}$ are i.i.d. sub-Gaussian random variables with equal probability from the set $\{\pm 1\}$.
- A2) Entries of $\mathbf{w}$ are i.i.d. Gaussian random variables with zero mean and variance σ^2.
- A3) $\mathbf{s}$ and $\mathbf{w}$ are independent.
- A4) $\mathbf{H}$ has full column rank.
- A5) All variables are real.

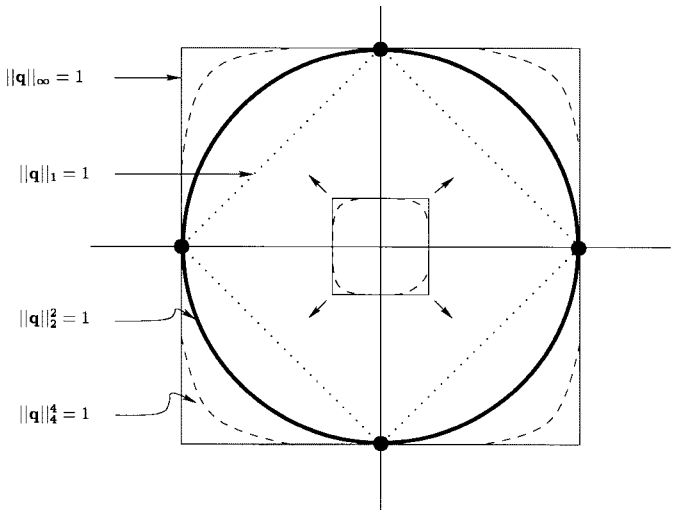

Fig. 2. Coincidence of various receivers when there is no noise.

Under A1)–A5) and the signal space property [20], CM, SW, and Wiener receivers can be obtained in the combined channel-receiver space $\mathbf{q}$ by optimizing the following cost functions [20]:

$$\text{CM:}\quad J_c(\mathbf{q}) \triangleq \overline{J}_c((\mathbf{H}^t)^\dagger\mathbf{q}) = 3\|\mathbf{q}\|_\mathbf{\Phi}^4 - 2\|\mathbf{q}\|_\mathbf{\Phi}^2 - 2\|\mathbf{q}\|_4^4 + 1 \tag{9}$$

$$\text{SW:}\quad J_s(\mathbf{q}) \triangleq \overline{J}_s((\mathbf{H}^t)^\dagger\mathbf{q}) = \|\mathbf{q}\|_4^4 \quad \text{subject to } \|\mathbf{q}\|_\mathbf{\Phi}^2 = 1 \tag{10}$$

$$\text{MSE:}\quad J_m^{(i)}(\mathbf{q}) \triangleq \overline{J}_m^{(i)}((\mathbf{H}^t)^\dagger\mathbf{q}) = \|\mathbf{q}\|_\mathbf{\Phi}^2 - 2\mathbf{q}^t\mathbf{e}_i + 1 \tag{11}$$

where

$$\mathbf{\Phi} = \mathbf{I} + \sigma^2(\mathbf{H}^t\mathbf{H})^{-1}. \tag{12}$$

A. Connections by Norms

To draw connections among various receivers, the key step in the geometrical approach is to convert the optimization of different cost functions to the same form. The one-one correspondence between SW and CM optima implies that CM and SW receivers can be obtained from the constrained optimization

$$\mathcal{O}_c\text{:}\ \min J_c(\mathbf{q}) \Longleftrightarrow \max_{\|\mathbf{q}\|_\mathbf{\Phi}^2=1} J_s(\mathbf{q}) \Longleftrightarrow \max_{\|\mathbf{q}\|_\mathbf{\Phi}^2=1} \|\mathbf{q}\|_4^4. \tag{13}$$

The equivalence "$\Longleftrightarrow$" above means that the optima of two optimizations are colinear and one-one correspondent. It is shown in Appendix A that Wiener receivers $\{\mathbf{q}_m^{(i)}\}$ are the results of the following maximization of the same type:

$$\mathcal{O}_m\text{:}\ \min J_m^{(i)}(\mathbf{q}), \quad i = 1, \cdots, n \Longleftrightarrow \max_{\|\mathbf{q}\|_\mathbf{\Phi}^2=1} \|\text{diag}^{-1}(\mathbf{\Phi}^{-1})\mathbf{q}\|_\infty \tag{14}$$

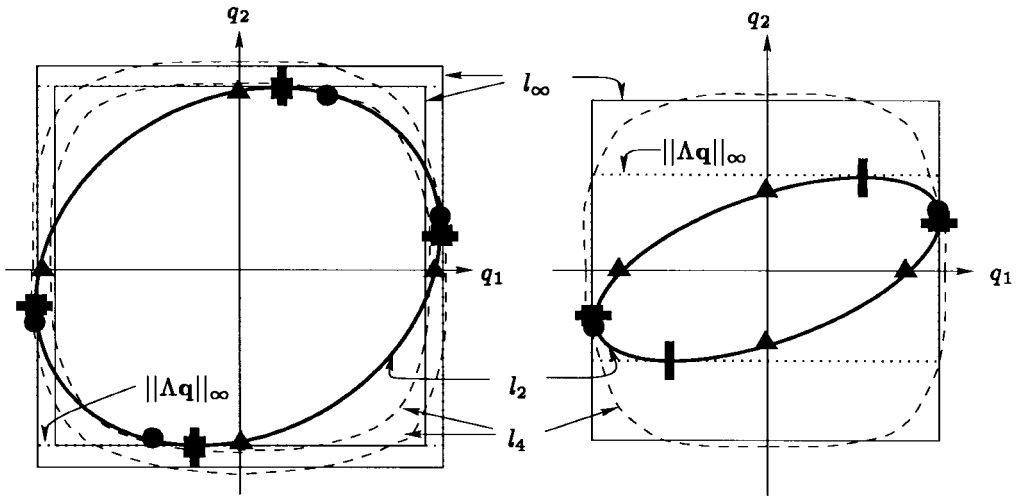

Fig. 3. (a) High SNR case with two CM/SW receivers and (b) low SNR case with only one CM/SW receiver. Rectangle: Wiener receiver. Circle: CM/SW receiver. Triangle: ZF receiver. Square: maximum of l_∞. $\Lambda = \text{diag}^{-1}(\mathbf{\Phi}^{-1})$.

where $\text{diag}(\mathbf{\Phi}^{-1}) = \text{diag}(\psi_{11}, \cdots, \psi_{nn})$ if $\mathbf{\Phi}^{-1} = [\psi_{ij}]_{i,j=1}^{n}$. We further consider the constrained l_∞ optimization

$$\mathcal{O}_\infty: \quad \max_{||\mathbf{q}||_{\mathbf{\Phi}}^2=1} ||\mathbf{q}||_\infty \tag{15}$$

which serves as a bridge between Wiener and CM/SW receivers because i) all maxima of $\mathcal{O}_\infty$ are Wiener receivers, and ii) the l_∞-norm can be considered as an approximation of the l_4-norm in $\mathcal{O}_c$, especially for low-dimensional cases. This geometrical approximation demonstrates the relationship between Wiener and CM/SW receivers. Note also that $\mathcal{O}_\infty$ is far simpler than $\mathcal{O}_c$ and has closed-form solutions.

B. Geometrical Interpretation

To gain insight into this approach, consider first the noiseless case when $\mathbf{\Phi} = \mathbf{I}$. As shown in Fig. 2, the optimizations in (13)–(15) are equivalent to expanding an l_4- (or l_∞-) norm ball until it is tangent outside of the unit ball. CM/SW (or Wiener) receivers are obtained at the tangent points. By modifying the linear constraint to the output power constraint $||\mathbf{q}||_{\mathbf{\Phi}}^2 = 1$, it is also interesting to consider the minimization of the convex cost function for VVK receivers

$$J_v(\mathbf{q}) \triangleq \overline{J}_v((\mathbf{H}^t)^\dagger \mathbf{q}) = ||\mathbf{q}||_1 \tag{16}$$

which is the result of shrinking the l_1-norm ball (the diamond) until its vertices touch the unit ball. This geometrical approach enables us to conclude that the blind receivers (CM, SW, and VVK) and the nonblind receivers (Wiener and ZF) are identical when there is no noise. The algebraic proofs of the same result [5], [14], [19] appear to be far more involved.

Situations in the presence of noise are more complicated. The power constraint $||\mathbf{q}||_{\mathbf{\Phi}}^2 = 1$ defines an ellipsoid instead of the unit ball. The shape of the ellipsoid depends on the noise level σ and the channel matrix $\mathbf{H}$. CM/SW receivers are again obtained by circumscribing an l_4-norm ball to the ellipsoid. For the 2-D case, we can observe two scenarios. Fig. 3(a) shows the case (occurring sometimes at high SNR) when there are two CM/SW receivers (only those receivers in the upper half plane are considered owing to the symmetry of all these norms), both are close to the corresponding Wiener receivers. Due to the geometrical similarity between l_∞- and l_4-norms, the same behavior can also be observed in the optimization of $\mathcal{O}_\infty$. Note that Wiener receivers are obtained from $\mathcal{O}_m$ by circumscribing a *rectangle* (weighted l_∞-norm ball) to the ellipsoid. For this (high SNR) case, the optima of $\mathcal{O}_m$ and $\mathcal{O}_\infty$ are identical. In other words, the connection between CM/SW and Wiener receivers is rooted in the connection among l_4-, l_∞-, and weighted l_∞-norms. Fig. 3(b) illustrates the situation (occurring sometimes at low SNR) when there exists only one CM/SW receiver (for s_1), which is also close to the Wiener receiver. The l_∞-norm optimization again behaves similarly. While the above two scenarios can be understood intuitively with the geometrical approach, it is not clear on what condition the CM/SW cost has one, two, or even more local optima. Later, in Section IV, a necessary and sufficient condition is presented to determine the number of CM/SW receivers for the 2-D case.

III. Results in High Dimensions

A. When Do CM, SW, and Wiener Receivers Coincide?

We have shown that if there is no noise, all receivers considered earlier coincide. It is also observed that in certain noisy case (shown in Fig. 3), CM/SW receivers are close to, but not the same as, Wiener receivers. The question is: *Are there any other cases that CM/SW receivers are colinear with, and one-one correspondent to, Wiener receivers?* One such scenario happens when the major axis of the ellipsoid coincides with $\mathbf{e}_i$, as illustrated in Fig. 4, which corresponds to the columns of $\mathbf{H}$ being orthogonal (but not necessarily having the same norm). In [2], this is also observed for the channels with orthonormal columns. We show next that there is no other possibility that CM/SW and Wiener receivers are equivalent.

Theorem 1: CM, SW, Wiener, and ZF receivers are colinear and one-one correspondent if and only if either there is no noise or the columns of $\mathbf{H}$ are orthogonal.

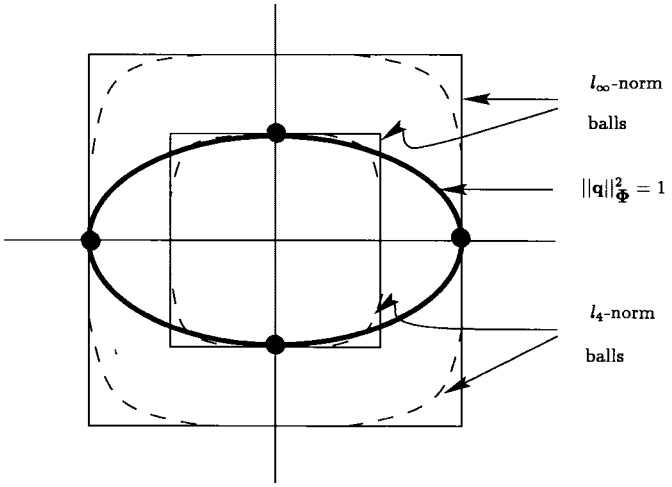

Fig. 4. Coincidence of CM/SW receivers with Wiener receivers when the columns of $\mathbf{H}$ are orthogonal.

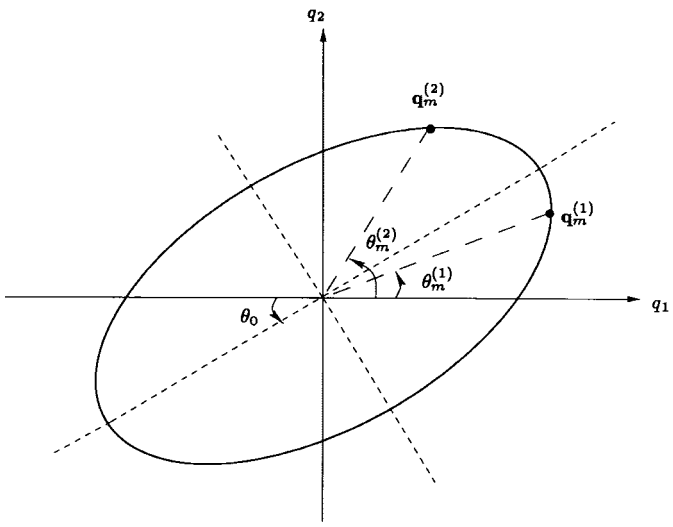

Fig. 5. Geometrical illustrations. θ_0 is the angle contained by the major axis of the ellipsoid and q_1 axis; $\mathbf{q}_m^{(i)}$ corresponds to the Wiener receiver for s_i; $\theta_m^{(i)}$ is the angle corresponding to $\mathbf{q}_m^{(i)}$.

Proof: See Appendix B.

The fact that blind CM/SW receivers can perform equally well as nonblind Wiener receivers under an arbitrary noise level has not been reported in the open literature. The orthogonal channel case considered in the above theorem plays an important part in the synchronous CDMA and OFDM systems. For example, in some CDMA applications, the channel matrix can be written as $\mathbf{H} = [\mathbf{c}_1, \cdots, \mathbf{c}_K]\,\mathrm{diag}\big(\sqrt{\mathcal{E}_1}, \cdots, \sqrt{\mathcal{E}_K}\big)$ with $\mathbf{c}_i$ and $\mathcal{E}_i$ being the spreading code and the power of the ith user, respectively, which is used to deal with the issue of the initialization of equalizers, i.e., the domains of attraction [7], [8].

B. Difficulties in the General Analysis

A full characterization of the constant modulus algorithm necessitates the analysis in high dimensions. Answers to Q1)–Q3) would result from the solution to the following constrained optimization problem:

$$\max_{||\mathbf{q}||^2_{\mathbf{\Phi}}=1} ||\mathbf{q}||_4^4. \tag{17}$$

Simple as it seems, this optimization turns out to be difficult for arbitrary positive definite $\mathbf{\Phi}$. We can take the Lagrangian approach, which again leads to numerical solutions that are local. In the interest of gaining insights into the behavior of CM/SW receivers and their relationship with Wiener receivers, we consider the 2-D case.

IV. Results in Two Dimensions

In this section, convergence properties of CM/SW receivers for the general 2-D channel-receiver impulse response are characterized. We formally establish that for every CM/SW receiver, there is a corresponding Wiener receiver nearby.

It is convenient to use the polar coordinate. Given a vector $\mathbf{q}$, let $\theta = \angle\mathbf{q}$, $\rho = ||\mathbf{q}||_2$, and then

$$\mathbf{q} = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} \triangleq \begin{pmatrix} \rho\cos(\theta) \\ \rho\sin(\theta) \end{pmatrix}. \tag{18}$$

Let $\mathbf{q}$ be an estimator of s_i. Define the signal-to-interference ratio by

$$\gamma_i(\mathbf{q}) \triangleq \frac{|q_i|}{\sum_{j,\, j\neq i} |q_j|}. \tag{19}$$

As illustrated in Fig. 5, the shape of the ellipsoid is determined by angle θ_0 of the major axis and the condition number $\chi(\mathbf{\Phi})$. Due to the symmetry of the CM and SW cost functions, without loss of generality, we only need to consider solutions for $\theta_0 \in (0, \pi/4]$ and the upper half of the ellipsoid, i.e., $\theta \in [0, \pi)$.

Characterizations of CM/SW Receivers: The following theorem completely characterizes the number and the locations of CM/SW receivers. Further, a necessary and sufficient condition is given in terms of SIR's of Wiener receivers.

Theorem 2: Let $\mathbf{q}_m^{(i)}$ be the Wiener receiver for s_i with SIR $\gamma_i \triangleq \gamma_i(\mathbf{q}_m^{(i)})$. The number and the locations of CM/SW receivers are characterized by γ_i as follows:

T1) There exists at least one, at most two, CM/SW receivers. Except when $\theta_0 = \pi/4$ and $\gamma_2 \leq 2$,[2] each CM/SW receiver is associated with a unique source. Specifically, for each $\mathbf{q}_c$, there exists a unique s_i such that $\gamma_i(\mathbf{q}_c) > \gamma_j(\mathbf{q}_c)$, $\forall\, j \neq i$.

T2) There are two CM/SW receivers if and only if

$$\frac{\sqrt{\gamma_1\gamma_2 - 1}}{\gamma_1^2} < \frac{\sqrt{3}}{4} \tag{20}$$

$$\frac{\gamma_2 - \gamma_d}{1 - \gamma_1\gamma_d} + \frac{1}{\gamma_d^3} < 0 \tag{21}$$

where

$$\gamma_d = \frac{\sqrt{3}\,\gamma_1 + \sqrt{3\gamma_1^2 - 4\sqrt{3}\sqrt{\gamma_1\gamma_2 - 1}}}{2\sqrt{\gamma_1\gamma_2 - 1}}. \tag{22}$$

[2]When $\theta_0 = \pi/4$ and $\gamma_2 \leq 2$, there exists only one CM/SW receiver at $\angle\mathbf{q}_c^{(1)} = \pi/4$. We consider this case as a zero probability event.

T2′) If the SIR of the Wiener receiver for the weak signal is less than or equal to 6 dB, i.e., $\gamma_2 \leq 2$, then there is only one CM/SW receiver.

T3) Let $\theta_c^{(i)} = \angle \mathbf{q}_c^{(i)}$ be the angle corresponding to the CM/SW estimator for s_i. Then

$$\theta_c^{(i)} \in \left(\theta_L^{(i)}, \theta_U^{(i)}\right) \subset \left(\angle \mathbf{q}_m^{(1)}, \angle \mathbf{q}_m^{(2)}\right) \tag{23}$$

where

$$\theta_L^{(1)} = \arctan\left(\frac{\gamma_1^3+\gamma_2}{1+\gamma_1^4}\right), \theta_U^{(1)} = \arctan\left(\frac{1+\gamma_2\gamma_e^3}{\gamma_1+\gamma_e^3}\right) \tag{24}$$

$$\theta_L^{(2)} = \arctan\left(\frac{1+\gamma_2\gamma_q^3}{\gamma_1+\gamma_q^3}\right), \theta_U^{(2)} = \arctan\left(\frac{1+\gamma_2^4}{\gamma_1+\gamma_2^3}\right) \tag{25}$$

$$\gamma_e = \min\left\{1, \frac{\sqrt{3}\gamma_1 - \sqrt{3\gamma_1^2 - 4\sqrt{3}\sqrt{\gamma_1\gamma_2 - 1}}}{2\sqrt{\gamma_1\gamma_2 - 1}}\right\} \tag{26}$$

$$\gamma_q = \frac{\sqrt{3}\gamma_2 + \sqrt{3\gamma_2^2 - 4\sqrt{3}\sqrt{\gamma_1\gamma_2 - 1}}}{2\sqrt{3}}. \tag{27}$$

T4) The SIR of a CM/SW receiver is less than that of the corresponding Wiener receiver, i.e., $\gamma_i(\mathbf{q}_c^{(i)}) < \gamma_i(\mathbf{q}_m^{(i)})$.

Proof: See Appendix C.

We now discuss some of the implications of Theorem 2. Our discussions are focused on two issues: the number of optima for the CM/SW cost and the locations of CM/SW receivers.

The Number of CM/SW Receivers: Because the surface of the ellipsoid is compact, $\mathcal{O}_c$ must have at least one local optimum. This is also observed in [2]. The fact that there are no more than two local optima implies that the CM/SW cost function does not create spurious local optima, i.e., each local optimum will correspond to one and only one estimated signal. Equation (23) indicates that they both are close to the Wiener receivers. However, it is possible that CM/SW receivers are not able to estimate a particular signal as observed in simulation [12]. T1) establishes this phenomenon formally. In other words, if a CM/SW receiver for s_2 is not in the neighborhood of the Wiener receiver, it simply does not exist. See Fig. 3(b).

It is then important to test when CM/SW receivers may fail to estimate one signal. T2′) shows that if the lower SIR of the two Wiener receivers is less than or equal to 6 dB, then s_2 cannot be estimated. Another way is to use the optimization involving the l_∞-norm, which also leads to Wiener receivers. It can be shown that if optimization $\mathcal{O}_\infty$ has only global maximum, then there is only one CM/SW receiver.[3] A more powerful test is given in T2), which depends only on SIR's of Wiener receivers. The generalization of T1) to the high-dimensional case is not trivial, although no counter example has been reported.

The number of CM/SW receivers also depends on the orientation (θ_0) of the ellipsoid. We have already established that when $\theta_0 = 0$, there are always two CM/SW receivers, regardless of the noise level and the channel condition. Interestingly, if $\theta_0 > 0$, then at low SNR, there always exists a channel with large enough condition number so that one CM/SW receiver disappears. This can be seen from T2′).

The Locations of CM/SW Receivers: It is interesting to note from (23) that the locations of CM/SW receivers are always within the sectors determined by Wiener receivers, which implies that the CM/SW receiver $\mathbf{q}_c^{(i)}$ has less signal component but more interference than the corresponding Wiener receiver $\mathbf{q}_m^{(i)}$. Consequently, the SIR of the CM/SW receiver is less than that of the Wiener receiver [T4)]. This relation also leads to a location bound for CM/SW receivers. A better bound, however, is given by $\theta_L^{(i)}$ and $\theta_U^{(i)}$. The accuracy of this bound is demonstrated in the numerical example presented next. In fact, the location bound can be made arbitrarily accurate by using the iterative technique in the proof of T3).

When SNR is low and the channel is ill-conditioned, it seems that CM/SW receivers should have worse SIR performance. It turns out that this may not be true. In fact, the SIR performance of the global CM/SW receiver approaches that of the Wiener receiver as the condition number of the channel matrix increases at low SNR. From the geometrical view, this counterintuitive phenomenon is not difficult to explain: the larger the condition number $\chi(\mathbf{\Phi})$, the larger the curvature of the ellipsoid. Therefore, the tangent points can only be obtained near the tips of the ellipsoid regardless of the type of norm balls. This gives the reason why CM/SW and Wiener receivers are close to each other. See also the numerical example shown next.

Numerical Examples: The locations and the performance of CM/SW receivers are affected by two factors that determine the shape of the ellipsoid: the orientation θ_0 and the condition number $\chi(\mathbf{\Phi})$. Instead of fixing the channel or SNR, we consider the combined effect of the channel and the noise by evaluating the locations of CM/SW receivers against θ_0 and $\chi(\mathbf{\Phi})$.

Fig. 6 illustrates the locations of CM/SW receivers versus the condition number for a fixed θ_0. The location bounds given in (23) along with the location of Wiener receivers are also plotted. When the condition number is small, the ellipsoid approximates the unit circle. It can be seen that CM/SW and Wiener receivers are close, and the location bounds are tight. As $\chi(\mathbf{\Phi})$ increases, the CM/SW receiver for s_2 starts to deviate from the Wiener receiver and eventually disappears as the condition number exceeds 2 [the exact value can be calculated from T2)]. On the other hand, the global optimum (for s_1) behaves differently. As $\chi(\mathbf{\Phi})$ increases, it departs from the corresponding Wiener receiver around $\chi(\mathbf{\Phi}) = 10$, and interestingly, they merge again for large $\chi(\mathbf{\Phi})$. As explained previously, for large $\chi(\mathbf{\Phi})$, the global CM/SW receiver and the Wiener receiver are close. In fact, the locations of both Wiener and CM/SW receivers approach θ_0. The largest discrepancy between them occurs when the ellipsoid is neither too "round" nor too "sharp" and θ_0 is close to $\pi/4$.

Fig. 7 shows the situation when the columns of $\mathbf{H}$ are nearly orthogonal (small θ_0). It is evident that the global CM/SW and the Wiener receivers are almost colinear throughout the

[3] Further, if there are two CM/SW receivers, there must be two maxima of $\mathcal{O}_\infty$.

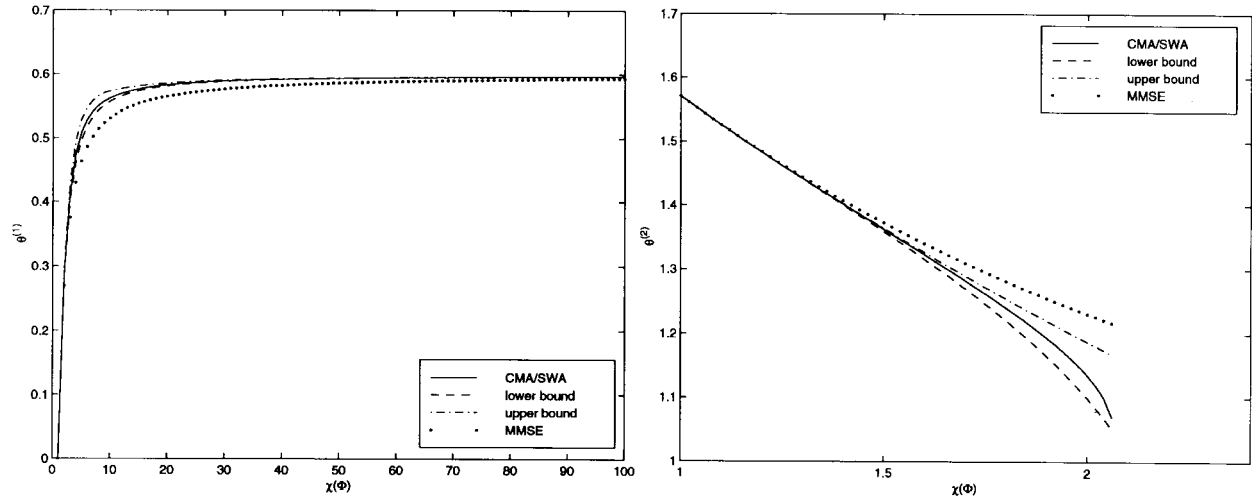

Fig. 6. Locations and location bounds of CM/SW receivers ($\theta_0 = 0.5992$). Left: Global optimum (for s_1). Right: Local optimum (for s_2).

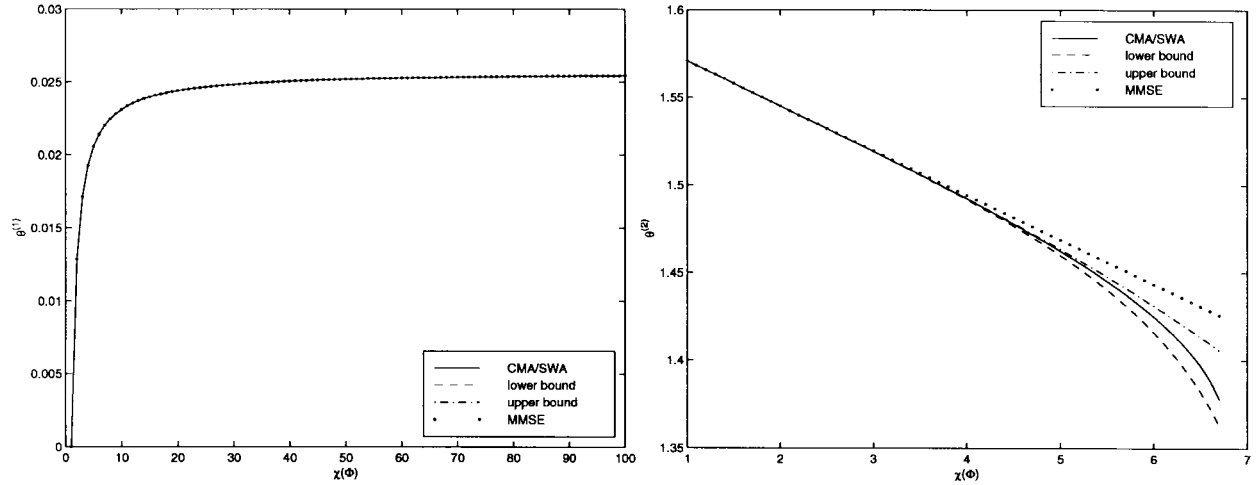

Fig. 7. Locations and location bounds of CM/SW receivers ($\theta_0 = 0.0257$). Left: Global optimum (for s_1). Right: Local optimum (for s_2).

entire range of the condition number $\chi(\mathbf{\Phi})$. However, the local CM/SW receiver separates from the Wiener receiver as $\chi(\mathbf{\Phi})$ increases and disappears at a larger $\chi(\mathbf{\Phi})$ compared with the previous figure.

V. Conclusion

In this paper, we presented a new geometrical approach to the analysis of CM and SW receivers in the presence of noise. By transforming various blind receiver design problems to expanding (or shrinking) norm balls of different types constrained on an ellipsoid, this approach reveals the connection among CM, SW, and Wiener receivers. This connection is rooted in the geometrical similarity among l_4-, l_∞-, and weighted l_∞-norm balls. For the important class of orthogonal channels, we are able to answer fundamental problems Q1)–Q3) completely.

Some of the conclusions appear to hold in the higher dimensions. It confirms that, although only for the 2-D case, CM/SW receivers do not have local minima that are not associated with Wiener receivers. However, they may fail to estimate some signals depending on the channel and noise conditions.

The analysis of CM/SW receivers for the 2-D case is significant: not because of its generality but for the insights into the behavior of CM/SW receivers. The conjecture that all CM receivers are associated with Wiener receivers can now be proved for the 2-D case. While the generalization of the 2-D results presented here to the general high-dimensional case, if successful, will allow us to quantitatively assess the performance loss of blind receivers when they are compared with (nonblind) Wiener receivers, it should be noted that the techniques employed in this paper to obtain results for the 2-D case do not have simple extension for the high-dimensional case. In order to fully understand the behavior of CM receivers, new techniques are needed. It is our hope that the geometrical approach presented in this paper will offer useful hints.

Appendix A
Proof of (14)

We first prove that

$$\min J_m^{(i)}(\mathbf{q}) \Longleftrightarrow \max_{\|\mathbf{q}\|_{\mathbf{\Phi}}^2=1} |\mathbf{q}^t \mathbf{e}_i|. \tag{28}$$

Let $\mathbf{q} = \alpha\mathbf{v}$, where α is a scalar, and $\|\mathbf{v}\|_{\mathbf{\Phi}}^2 = 1$. The MSE cost function in (11) becomes

$$J_m^{(i)}(\alpha\mathbf{v}) = \alpha^2 - 2\alpha\mathbf{v}^t\mathbf{e}_i + 1. \tag{29}$$

Then, we have

$$\begin{aligned}\min J_m^{(i)}(\mathbf{q}) &\Leftrightarrow \min_{\|\mathbf{v}\|_{\mathbf{\Phi}}^2=1} \min_{\alpha} \alpha^2 - 2\alpha\mathbf{v}^t\mathbf{e}_i + 1 \\ &= \min_{\|\mathbf{v}\|_{\mathbf{\Phi}}^2=1} 1 - (\mathbf{v}^t\mathbf{e}_i)^2 \\ &\Leftrightarrow \max_{\|\mathbf{v}\|_{\mathbf{\Phi}}^2=1} |\mathbf{v}^t\mathbf{e}_i| \Leftrightarrow \max_{\|\mathbf{q}\|_{\mathbf{\Phi}}^2=1} |\mathbf{q}^t\mathbf{e}_i|. \end{aligned} \tag{30}$$

Next, let $\text{diag}(\mathbf{\Phi}^{-1}) = \text{diag}(\psi_{11}, \cdots, \psi_{nn})$. Then

$$\begin{aligned}&\max_{\|\mathbf{q}\|_{\mathbf{\Phi}}^2=1} \|\text{diag}^{-1}(\mathbf{\Phi}^{-1})\mathbf{q}\|_\infty \\ &= \max_{\|\mathbf{q}\|_{\mathbf{\Phi}}^2=1} \max_i \left|\frac{q_i}{\psi_{ii}}\right| \\ &= \max_i \max_{\|\mathbf{q}\|_{\mathbf{\Phi}}^2=1} \left|\frac{\mathbf{q}^t\mathbf{e}_i}{\psi_{ii}}\right| \\ &= \max_i \frac{1}{|\psi_{ii}|} \max_{\|\mathbf{q}\|_{\mathbf{\Phi}}^2=1} |\mathbf{q}^t\mathbf{e}_i|. \end{aligned} \tag{31}$$

Equation (28) implies that the Wiener receiver $\mathbf{q}_m^{(i)} = \mathbf{\Phi}^{-1}\mathbf{e}_i$ is the optimal solution to (31), which results in

$$\begin{aligned}\max_{\|\mathbf{q}\|_{\mathbf{\Phi}}^2=1} \|\text{diag}^{-1}(\mathbf{\Phi}^{-1})\mathbf{q}\|_\infty &= \max_i \frac{1}{|\psi_{ii}|} |(\mathbf{\Phi}^{-1}\mathbf{e}_i)^t\mathbf{e}_i| \\ &= \max_i \frac{1}{|\psi_{ii}|} |\psi_{ii}| = 1. \end{aligned} \tag{32}$$

This proves the equivalent relation in (14). □□□

Appendix B
Proof of Theorem 1

The techniques adopted here involve the Lagrangian algorithm and the concept of tangent planes [1], [13].

1) Sufficiency: The two cases correspond to diagonal $\mathbf{\Phi}$, i.e., $\mathbf{\Phi} = \text{diag}(\phi_{11}, \cdots, \phi_{nn})$. First, the Wiener and the ZF receivers for signal s_i are equivalent, i.e., they are one-one correspondent and colinear:

$$\mathbf{q}_m^{(i)} = \mathbf{\Phi}^{-1}\mathbf{e}_i = \frac{1}{\phi_{ii}}\mathbf{e}_i \tag{33}$$

$$\mathbf{q}_z^{(i)} = \frac{1}{\sqrt{\phi_{ii}}}\mathbf{e}_i = \sqrt{\phi_{ii}}\,\mathbf{q}_m^{(i)}. \tag{34}$$

Second, it has been shown in [10] that SW and CM receivers are equivalent. We now show that SW and ZF receivers are equivalent.

Consider the constrained optimization for CM/SW receivers

$$\max_{h(\mathbf{q})=0} \|\mathbf{q}\|_4^4, \qquad h(\mathbf{q}) = \|\mathbf{q}\|_{\mathbf{\Phi}}^2 - 1. \tag{35}$$

The Lagrangian function is defined by

$$L(\mathbf{q}, \lambda) = \sum_{i=1}^n q_i^4 + \lambda\left(\sum_{i=1}^n \phi_{ii}q_i^2 - 1\right) \tag{36}$$

with gradient

$$\nabla L(\mathbf{q}, \lambda) = \begin{pmatrix} 4q_1^3 + 2\lambda q_1\phi_{11} \\ \vdots \\ 4q_n^3 + 2\lambda q_n\phi_{nn} \end{pmatrix}. \tag{37}$$

The stationary point $\mathbf{q}^*$ and λ^* can be solved from $\nabla L(\mathbf{q}, \lambda) = \mathbf{0}$. Furthermore, the first-order feasible variation yields

$$\nabla h^t(\mathbf{q}^*)\mathbf{y} = q_1^*\phi_{11}y_1 + \cdots + q_n^*\phi_{nn}y_n = 0. \tag{38}$$

The Hessian at the stationary point $\mathbf{q}^*$ is given by

$$\nabla^2 L(\mathbf{q}^*, \lambda^*) = \text{diag}(12q_1^{*2} + 2\lambda^*\phi_{11}, \cdots, 12q_n^{*2} + 2\lambda^*\phi_{nn}). \tag{39}$$

Without loss of generality, we assume $q_1^*, \cdots, q_k^* \neq 0$, $q_{k+1}^*, \cdots, q_n^* = 0$. Then, from (37) and $\nabla L(\mathbf{q}^*, \lambda^*) = \mathbf{0}$, (39) becomes

$$\begin{aligned}&\nabla^2 L(\mathbf{q}^*, \lambda^*) \\ &= \text{diag}(8q_1^{*2}, \cdots, 8q_k^{*2}, 2\lambda^*\phi_{k+1,k+1}, \cdots, 2\lambda^*\phi_{nn}) \end{aligned} \tag{40}$$

where λ^* is negative. If $k = n$, i.e., $q_i^* \neq 0\ \forall i$, then $\mathbf{q}^*$ is a minimum due to $\nabla^2 L(\mathbf{q}^*, \lambda^*) > 0$. If $k = 1$, which corresponds to the zero forcing solution, we have, for $\forall\mathbf{y}$ such that $\nabla h^t(\mathbf{q}^*)\mathbf{y} = 0$, $y_1 = 0$, and

$$\mathbf{y}^t\nabla^2 L(\mathbf{q}^*, \lambda^*)\mathbf{y} < 0. \tag{41}$$

Therefore, $\mathbf{q}^* = \sqrt{-(\lambda^*/2)\,\phi_{11}}\mathbf{e}_1$ is a local maximum that is colinear with and one-one correspondent to the Wiener receiver $\mathbf{q}_m^{(1)}$. For other k, $\mathbf{q}^*$ is a saddle point.

2) Necessity: We need to show that if the maxima of $\mathcal{O}_c$ have the form $\mathbf{q}_c^{(k)} = q_k\mathbf{e}_k$, $q_k \neq 0$, then $\mathbf{\Phi} = \text{diag}(\phi_{11}, \cdots, \phi_{nn})$. The Lagrangian function is defined by

$$L(\mathbf{q}, \lambda) = \sum_{i=1}^n q_i^4 + \lambda\left(\sum_{i=1}^n \sum_{j=1}^n \phi_{ij}q_iq_j - 1\right) \tag{42}$$

with gradient

$$\nabla L(\mathbf{q}, \lambda) = \begin{pmatrix} 4q_1^3 + 2\lambda(\phi_{11}q_1 + \cdots + \phi_{1n}q_n) \\ \vdots \\ 4q_n^3 + 2\lambda(\phi_{n1}q_1 + \cdots + \phi_{nn}q_n) \end{pmatrix}. \tag{43}$$

Then

$$\nabla L(\mathbf{q}_c^{(k)}, \lambda^*) = \begin{pmatrix} 2\lambda^*\phi_{1k}q_k \\ \vdots \\ 4q_k^3 + 2\lambda^*\phi_{kk}q_k \\ \vdots \\ 2\lambda^*\phi_{nk}q_k \end{pmatrix} = \mathbf{0}. \tag{44}$$

Since $q_k \neq 0$, $\lambda^* \neq 0$. Hence, $\phi_{ik} = 0$ for $i \neq k$, which implies that $\mathbf{\Phi}$ is diagonal. □□□

Appendix C
Proof of Theorem 2

Since CM/SW receivers are obtained from the constrained optimization $\mathcal{O}_c$ (13), we now prove this theorem by maximizing the following cost function:

$$h_4(\mathbf{q}), \qquad \text{subject to } h_2(\mathbf{q}) = 0 \tag{45}$$

where $h_4(\mathbf{q}) \triangleq \|\mathbf{q}\|_4^4$, $h_2(\mathbf{q}) \triangleq \|\mathbf{q}\|_{\mathbf{\Phi}}^2 - 1$, and $\mathbf{\Phi} \triangleq \left(\begin{smallmatrix}\phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22}\end{smallmatrix}\right)$.

When $\|\mathbf{q}\|_4^4$ achieves its maximum, the l_4-norm ball must be tangent outside of the ellipsoid, i.e., at the maximal point, the two norm balls have the same gradient:

$$\nabla h_2(\mathbf{q}) = \nabla h_4(\mathbf{q}). \tag{46}$$

The roots of (46) determine the properties of CM/SW receivers. For convenience, we will use the polar coordinate representation (18) and denote $z \triangleq \tan(\angle \mathbf{q})$, $z_1 \triangleq 1/\gamma_1$, $z_2 \triangleq \gamma_2$. Solving the optimization $\mathcal{O}_m$ (14), we obtain $z_1 = -(\phi_{12}/\phi_{22})$ and $z_2 = -(\phi_{11}/\phi_{12})$ by using (19). Without loss of generality, we only need to consider the case of $\chi(\mathbf{\Phi}) \neq 1$, $\theta \in [0, \pi)$, and $\theta_0 \in (0, \pi/4]$. Then, $\phi_{12} = \phi_{21} < 0$, $0 < \phi_{11} \leq \phi_{22}$. In this situation, we have $0 < z_1 < 1$, $z_1 z_2 \leq 1$, and $z_1 < z_2$.

We first show that solving (46) is equivalent to solving a fourth-order polynomial equation.

Lemma 1: Define

$$f(z) \triangleq \frac{z_1(z_2 - z)}{z_1 - z} + \frac{1}{z^3}. \tag{47}$$

$\nabla h_2(\mathbf{q}^*) = \nabla h_4(\mathbf{q}^*)$ if and only if $f(z^*) = 0$, where $z^* = q_2^*/q_1^*$.

Proof:

$$\begin{aligned}\nabla h_2(\mathbf{q}^*) = \nabla h_4(\mathbf{q}^*) &\Leftrightarrow \begin{pmatrix} 4q_1^{*3} \\ 4q_2^{*3} \end{pmatrix} = \begin{pmatrix} 2\phi_{11}q_1^* + 2\phi_{12}q_2^* \\ 2\phi_{22}q_2^* + 2\phi_{12}q_1^* \end{pmatrix} \\ &\Leftrightarrow \frac{1}{z^{*3}} = \frac{\phi_{11} + \phi_{12}z^*}{\phi_{22}z^* + \phi_{12}} \\ &\Leftrightarrow \frac{1}{z^{*3}} = \frac{z_1(z^* - z_2)}{z_1 - z^*} \\ &\Leftrightarrow f(z^*) = 0. \end{aligned} \tag{48}$$

□

Next, we consider properties of $f(z)$ and $f'(z)$ in different regions. Define

$$\begin{aligned} \mathcal{R}_1 &\triangleq (-\infty, 0), & \mathcal{R}_2 &\triangleq (0, z_1) \\ \mathcal{R}_3 &\triangleq (z_1, z_2), & \mathcal{R}_4 &\triangleq [z_2, \infty). \end{aligned} \tag{49}$$

$z = 0$ and $z = z_1$ are excluded because they are not roots of $f(z) = 0$.

Lemma 2: $f(z)$ and $f'(z)$ have the following properties:

1) Let $\Delta_- = 3 - 4\sqrt{3}\, z_1 \sqrt{z_1(z_2 - z_1)}$. $f'(z)$ have four zeros $\overline{z}_a \in \mathcal{R}_1$, $\overline{z}_b \in \mathcal{R}_2$, $\overline{z}_c$, $\overline{z}_d \in \mathcal{R}_3 \bigcup \mathcal{R}_4$ if and only if $\Delta_- \geq 0$; otherwise, $f'(z)$ has only two zeros $\overline{z}_a$ and $\overline{z}_b$. Further, if $\Delta_- > 0$, then $f(\overline{z}_c) > 0$.
2) a) In $\mathcal{R}_1$, there exists a unique z_a such that $f(z_a) = 0$ and $f'(z_a) < 0$.
   b) In $\mathcal{R}_2 \bigcup \mathcal{R}_4$, $f(z) > 0$.
   c) In $\mathcal{R}_3$, $f(z)$ has either: i) a unique zero z_b and $f'(z_b) \geq 0$ with "=" holding at $z_b = 1$, $z_1 = \frac{1}{2}$, and $z_2 = 2$ or ii) three zeros z_b, z_c, and z_d such that $z_b < z_c < z_d$, $f'(z_b) > 0$, $f'(z_c) < 0$, and $f'(z_d) > 0$.

Proof: Let

$$g_+(z) \triangleq \sqrt{z_1(z_2 - z_1)}\, z^2 + \sqrt{3}(z - z_1) \tag{50}$$

$$g_-(z) \triangleq \sqrt{z_1(z_2 - z_1)}\, z^2 - \sqrt{3}(z - z_1). \tag{51}$$

Taking the derivative of $f(z)$ (47), we have

$$f'(z) = \frac{g_+(z)g_-(z)}{(z_1 - z)^2 z^4} = \frac{(z - \overline{z}_a)(z - \overline{z}_b)(z - \overline{z}_c)(z - \overline{z}_d)}{(z_1 - z)^2 z^4} \tag{52}$$

where

$$\overline{z}_a = \frac{-\sqrt{3} - \sqrt{\Delta_+}}{2\sqrt{z_1(z_2 - z_1)}}; \qquad \overline{z}_b = \frac{-\sqrt{3} + \sqrt{\Delta_+}}{2\sqrt{z_1(z_2 - z_1)}} \tag{53}$$

$$\overline{z}_c = \frac{\sqrt{3} - \sqrt{\Delta_-}}{2\sqrt{z_1(z_2 - z_1)}}; \qquad \overline{z}_d = \frac{\sqrt{3} + \sqrt{\Delta_-}}{2\sqrt{z_1(z_2 - z_1)}} \tag{54}$$

$$\begin{aligned} \Delta_+ &= 3 + 4\sqrt{3}\, z_1 \sqrt{z_1(z_2 - z_1)} \\ \Delta_- &= 3 - 4\sqrt{3}\, z_1 \sqrt{z_1(z_2 - z_1)}. \end{aligned} \tag{55}$$

Since $\Delta_+ > 0$, $f'(z)$ can have four zeros if and only if $\Delta_- \geq 0$. The relationship among them satisfies

$$\overline{z}_a < 0 < \overline{z}_b < z_1 < \overline{z}_c \leq \overline{z}_d \tag{56}$$

if $\overline{z}_c$ and $\overline{z}_d$ exist. We next consider properties of $f(z)$ in the above defined regions.

In $\mathcal{R}_1$: This region is illustrated in Fig. 8. From (52), we have $f'(z) > 0$ in $(-\infty, \overline{z}_a)$ and $f'(z) < 0$ in $(\overline{z}_a, 0)$. Since $f(-\infty) > 0$ and $f(0^-) < 0$, $f(z)$ has a unique zero $z_a \in (\overline{z}_a, 0)$, and $f'(z_a) < 0$.

In $\mathcal{R}_2 \bigcup \mathcal{R}_4$: See Fig. 8. From (47), we can see that $f(z) > 0$. Therefore, no zero exists.

In $\mathcal{R}_3$: We need to consider several cases (Fig. 9). Note first that from (47), $f(z_1^+) < 0$, $f(1) \geq 0$, and $f(z_2) > 0$. Because $f'(z)$ has no more than two zeros ($\overline{z}_c$ and $\overline{z}_d$), $f(z) = 0$ can have at most three roots.

Case 1—$\Delta_- < 0$: Real $\overline{z}_c$ and $\overline{z}_d$ do not exist, and $f'(z) > 0$. There is a unique $z_b \in (z_1, 1]$ such that $f(z_b) = 0$ and $f'(z_b) > 0$ [Fig. 9(a)].

Case 2—$\Delta_- = 0$: In this situation, $\overline{z}_c = \overline{z}_d = 2z_1$, and $f'(z) \geq 0$ with "=" holding at $z = 2z_1$. Therefore, $f(z)$ has a unique zero z_b. We need to determine the property of $f(z_b)$ and $f'(z_b)$. $\Delta_- = 0$ leads to $z_2 = z_1 + (3/16z_1^3)$. Considering $z_1 z_2 \leq 1$, we have $z_1 \in [1/2, \sqrt{3}/2]$. Then, $f(\overline{z}_c) = f(2z_1) = z_1 - (1/16z_1^3) \geq 0$ with "=" satisfied at $z_1 = \frac{1}{2}$. Hence, $z_b \in (z_1, \overline{z}_c]$. If $z_b \neq \overline{z}_c$, $f(\overline{z}_c) > 0$, and $f'(z_b) > 0$. If $z_b = \overline{z}_c$, we have $f(\overline{z}_c) = f(z_b) = f'(z_b) = 0$, $z_b = 1$, $z_2 = 4z_1 = 2$; see Fig. 9(b).

Case 3—$\Delta_- > 0$: $f'(z) = 0$ has two different roots $\overline{z}_c$ and $\overline{z}_d$. If $\overline{z}_c \geq z_2$ [Fig. 9(c)], both $\overline{z}_c$ and $\overline{z}_d$ are in $\mathcal{R}_4$. Therefore, $f(\overline{z}_c) > 0$, $f(\overline{z}_d) > 0$, and $f(z)$ has only one zero z_b with $f'(z_b) > 0$. If $\overline{z}_c < z_2 \leq \overline{z}_d$ [Fig. 9(d)], then $f(\overline{z}_d) > 0$, $f'(z) > 0$ in $(z_1, \overline{z}_c)$, and $f'(z) < 0$ in $(\overline{z}_c, z_2)$. Only one zero $z_b \in (z_1, \overline{z}_c)$ exists such that $f'(z_b) > 0$ and $f(\overline{z}_c) > 0$. If $z_1 < \overline{z}_c < \overline{z}_d < z_2$, we need to prove $f(\overline{z}_c) > 0$. Recalling that $g_-(\overline{z}_c) = 0$ and $g_-(\overline{z}_d) = 0$ in (52), we have $\overline{z}_c^2 = \sqrt{3}(\overline{z}_c - z_1)/\sqrt{z_1(z_2 - z_1)}$, and $\overline{z}_d^2 = \sqrt{3}(\overline{z}_d - z_1)/\sqrt{z_1(z_2 - z_1)}$. Substituting them in (47) yields $f(\overline{z}_c) = [1/\sqrt{3}\,\overline{z}_c(\overline{z}_c - z_1)]u(\overline{z}_c)$, and $f(\overline{z}_d) = [1/\sqrt{3}\overline{z}_d(\overline{z}_d - z_1)]u(\overline{z}_d)$, where $u(z) = \sqrt{z_1(z_2 - z_1)} - \sqrt{3}\, z_1 z_2 z + \sqrt{3}\, z_1 z^2$. The zeros of $u(z)$ are

$$z_p = \frac{\sqrt{3}\, z_1 z_2 - \sqrt{\Delta_u}}{2\sqrt{3}\, z_1}, \qquad z_q = \frac{\sqrt{3}\, z_1 z_2 + \sqrt{\Delta_u}}{2\sqrt{3}\, z_1} \tag{57}$$

where $\Delta_u = 3z_1^2 z_2^2 - 4\sqrt{3}\, z_1 \sqrt{z_1(z_2 - z_1)}$. We now have three situations.

a) If $\Delta_u < 0$, then $u(z) > 0$, $f(\overline{z}_c) > 0$, $f(\overline{z}_d) > 0$. There is a unique zero z_b with $f'(z_b) > 0$. See Fig. 9(e).

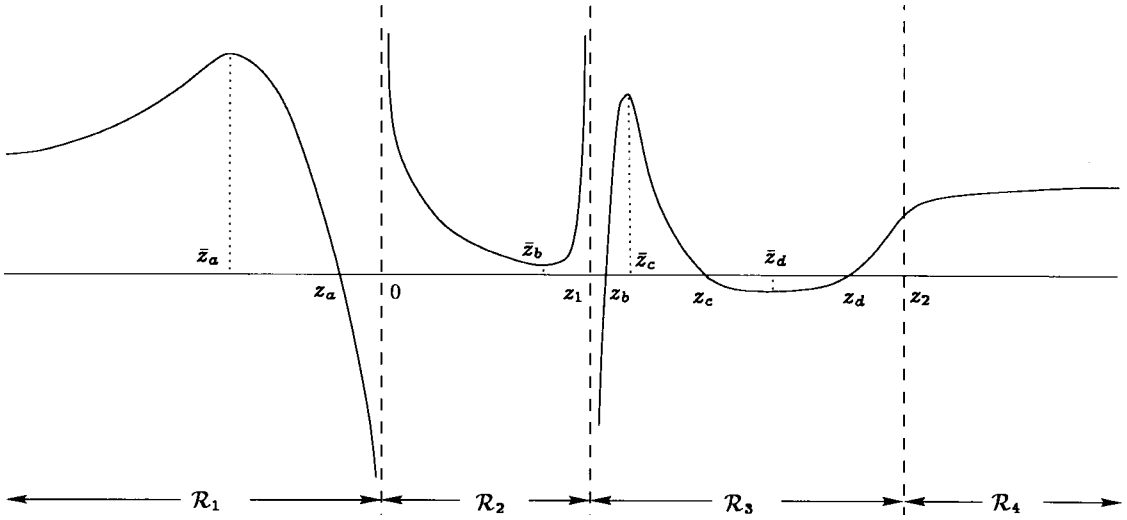

Fig. 8. Illustration of $f(z)$ in different regions.

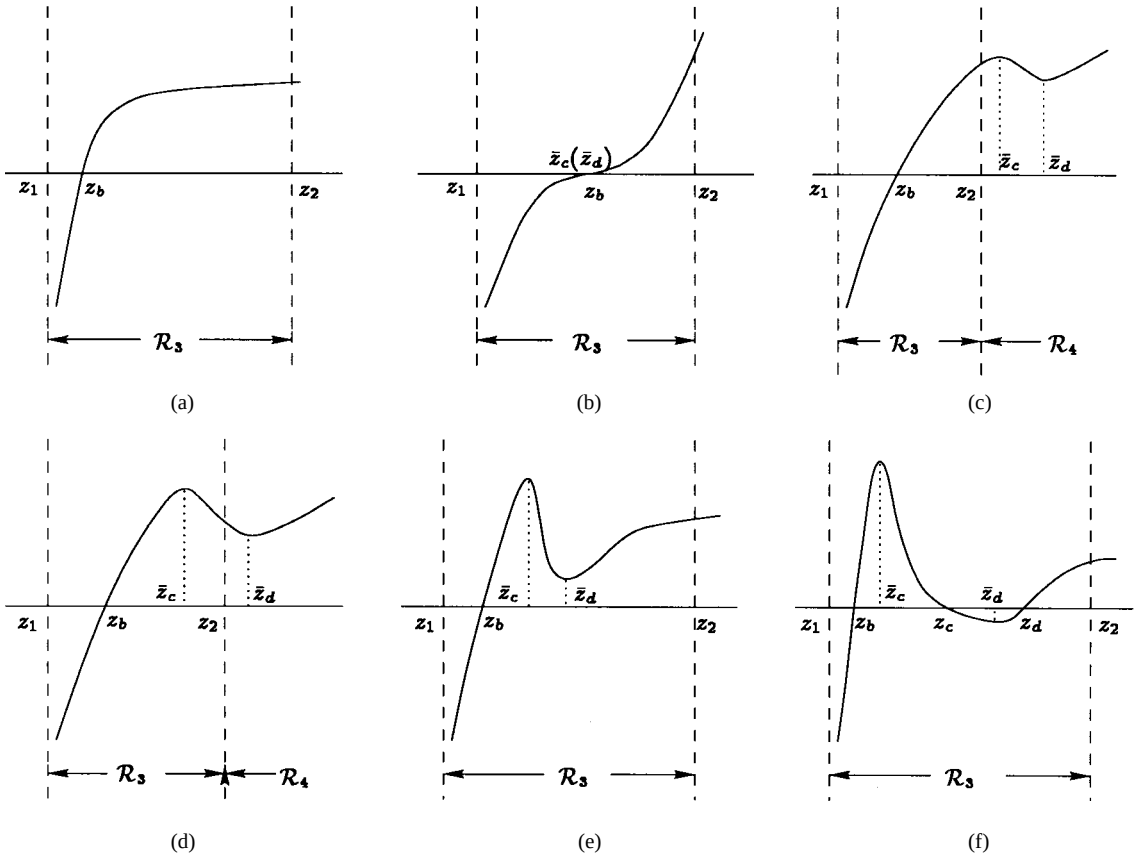

Fig. 9. Different behavior of $f(z)$ in $\mathcal{R}_3$.

b) If $\Delta_u = 0$, we have $f(\overline{z}_c) \geq 0$, $z_p = z_q = z_2/2$, and $\overline{z}_c = 2(1 - \sqrt{1 - z_1^2 z_2^2})/z_1 z_2^2$. When $\overline{z}_c \neq z_2/2$, the same conclusion as in a) can be drawn. If $\overline{z}_c = z_2/2$, then $f(\overline{z}_c) = f(\overline{z}_d) = 0$. We have $z_p = z_q = \overline{z}_c = \overline{z}_d = 1$, which contradicts the assumption $\overline{z}_c < \overline{z}_d$.

c) If $\Delta_u > 0$, considering $z_1 z_2 \leq 1$, we have $z_2^2 > 4$ leading to $z_2 > 4z_1$. Then, $\overline{z}_c < z_p$ can be obtained, and therefore, $f(\overline{z}_c) > 0$. Further, if $f(\overline{z}_d) < 0$, there exist three roots of $f(z) = 0$: $z_b \in (z_1, \overline{z}_c)$, $z_c \in (\overline{z}_c, \overline{z}_d)$, $z_d \in (\overline{z}_d, z_2)$ with $f'(z_b) > 0$, $f'(z_c) < 0$, and $f'(z_d) > 0$ [Fig. 9(f)]. Note that this is the only case where $f(\overline{z}_d) < 0$.

From the above analysis, we can conclude that $f(\overline{z}_c) > 0$ if $\Delta_- > 0$, and in $\mathcal{R}_3$, $f(z)$ can have either only one zero z_b and $f'(z_b) \geq 0$ with "=" holding at $z_b = 1$, $z_2 = 4z_1 = 2$ or three zeros with properties described in c). □

In the following Lemma, we provide a test (of z) for the optima of (45).

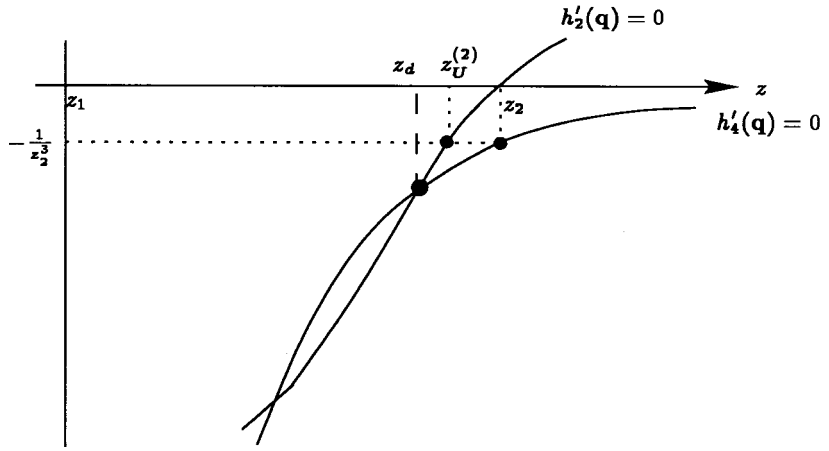

Fig. 10. Illustration of finding a tighter upper bound for the local CM/SW optimum.

Lemma 3: $\mathbf{q}^*$ is a stationary point of (45) if and only if $f(z^*)=0$. Further, $\mathbf{q}^*$ is a maximum (minimum) if and only if $f'(z^*)\geq 0$ ($f'(z^*)<0$).

Proof: For the maximization problem (45), the Lagrangian function is defined by

$$L(\mathbf{q},\,\lambda)=q_1^4+q_2^4+\lambda(\phi_{11}q_1^2+\phi_{22}q_2^2+2\phi_{12}q_1q_2-1) \quad (58)$$

with gradient

$$\begin{aligned}\nabla L(\mathbf{q},\,\lambda)&=\begin{pmatrix}4q_1^3\\4q_2^3\end{pmatrix}+\lambda\begin{pmatrix}2\phi_{11}q_1+2\phi_{12}q_2\\2\phi_{22}q_2+2\phi_{12}q_1\end{pmatrix}\\&=\nabla h_4(\mathbf{q})+\lambda\nabla h_2(\mathbf{q}).\end{aligned} \quad (59)$$

Then

$$\nabla L(\mathbf{q}^*,\lambda^*)=\mathbf{0}\Leftrightarrow\nabla h_2(\mathbf{q}^*)=\nabla h_4(\mathbf{q}^*)\Leftrightarrow f(z^*)=0 \quad (60)$$

which proves the first part of the Lemma.

By solving $\nabla L(\mathbf{q}^*,\,\lambda^*)=\mathbf{0}$, we have

$$\lambda^*=-2(q_1^{*4}+q_2^{*4}),\qquad \frac{\phi_{11}+\phi_{12}z^*}{\phi_{22}z^*+\phi_{12}}=\frac{1}{z^{*3}}. \quad (61)$$

Considering the first-order feasible variation

$$\nabla h_2^t(\mathbf{q}^*)\mathbf{y}=0,\qquad \mathbf{y}=(y_1\;y_2)^t \quad (62)$$

yields

$$\frac{y_2}{y_1}=-\frac{\phi_{11}+\phi_{12}z^*}{\phi_{22}z^*+\phi_{12}}. \quad (63)$$

The Hessian at the stationary point is given by

$$\nabla^2 L(\mathbf{q}^*,\lambda^*)=\begin{pmatrix}12q_1^{*2}&0\\0&12q_2^{*2}\end{pmatrix}+2\lambda^*\begin{pmatrix}\phi_{11}&\phi_{12}\\\phi_{12}&\phi_{22}\end{pmatrix}. \quad (64)$$

Let $r=y_2/y_1$. We have

$$\begin{aligned}&\mathbf{y}^t\nabla^2 L(\mathbf{q}^*,\,\lambda^*)\mathbf{y}\\&\quad=y_1^2(1\;r)\nabla^2 L(\mathbf{q}^*,\,\lambda^*)(1\;r)^t\\&\quad=2y_1^2[6q_1^{*2}+6r^2q_2^{*2}+(\phi_{11}+\phi_{22}r^2+2\phi_{12}r)\lambda^*].\end{aligned} \quad (65)$$

Substituting with (61), (63), and $\mathbf{y}^t\mathbf{\Phi}\mathbf{y}=1$ leads to

$$\begin{aligned}\mathbf{y}^t\nabla^2 L(\mathbf{q}^*,\,\lambda^*)\mathbf{y}&=\frac{4y_1^2(q_1^{*4}+q_2^{*4})}{q_1^{*2}}\left[\frac{3}{z^{*4}}-\frac{z_1(z_2-z_1)}{(z_1-z^*)^2}\right]\\&=-\beta f'(z^*)\end{aligned} \quad (66)$$

where β is positive. Therefore, when $f'(z^*)>0$, $\mathbf{y}^t\nabla^2 L\mathbf{y}<0$, z^* is a maximum; when $f'(z^*)<0$, $\mathbf{y}^t\nabla^2 L\mathbf{y}>0$, z^* is a minimum. From Lemma 2, $f(z^*)=f'(z^*)=0$ happens if and only if $z^*=1$, $z_1=\frac{1}{2}$, $z_2=2$. Substituting z_1 and z_2 in $\|\mathbf{q}\|_{\mathbf{\Phi}}^2=1$ yields $q_1^2+q_2^2-q_1q_2=c$ or $q_1^2(1-z+z^2)=c$, where c is a scalar. The cost (45) becomes $q_1^4+q_2^4=q_1^4(1+z^4)=c^2(1+z^4)/(z^2-z+1)^2\leq c^2$ with "=" satisfied at $z^*=1$. This proves that $z^*=1$ is still a maximum. The necessity of the second part is immediate from Lemma 1 and (66). □

We are now ready to prove Theorem 2.

Proof of T1): Using Lemmas 2 and 3, we conclude that in $\mathcal{R}_1$, there is a unique minimum of the cost (45), and there is no extremum in $\mathcal{R}_2\bigcup\mathcal{R}_4$. Applying Lemma 3 to $\mathcal{R}_3$, we can see that there is at least one at most two maxima. If $\theta_0\in(0,\,\pi/4)$, we have $f(1)>0$. Then, $\angle\mathbf{q}_c^{(1)}<\pi/4$, which implies $\gamma_1(\mathbf{q}_c^{(1)})>\gamma_2(\mathbf{q}_c^{(1)})$. If $\mathbf{q}_c^{(2)}$ exists, from T2′), we immediately have $\gamma_2(\mathbf{q}_c^{(2)})>\gamma_1(\mathbf{q}_c^{(2)})$. If $\theta_0=\pi/4$, then $z_1z_2=1$, $\gamma_1=\gamma_2$. We can solve $f(z)=0$ directly. If $\gamma_2>2$ ($z_1<\frac{1}{2}$), then $z_b=(1-\sqrt{1-4z_1^2})/2z_1$, $z_c=1$, and $z_d=(1+\sqrt{1-4z_1^2})/2z_1$. Therefore, two CM/SW receivers exist with $\gamma_1(\mathbf{q}_c^{(1)})>\gamma_2(\mathbf{q}_c^{(1)})$ and $\gamma_2(\mathbf{q}_c^{(2)})>\gamma_1(\mathbf{q}_c^{(2)})$. If $\gamma_2\leq 2$, there exists only one CM/SW receiver at $\angle\mathbf{q}_c^{(1)}=\pi/4$. In this case, $\gamma_1(\mathbf{q}_c^{(1)})=\gamma_2(\mathbf{q}_c^{(1)})$.

Proof of T2): To show the "if" part, we assume that (20) and (21) are satisfied. When (20) is true, $\Delta_-\,>0$. Therefore, $f'(z)$ has two different zeros $\overline{z}_c<\overline{z}_d$ [defined in (54)] and $f(\overline{z}_c)>0$ (Lemma 2). If (21) also holds, i.e., $f(\overline{z}_d)<0$, then (45) has two maxima.

To show the "only if" part, we assume two CM/SW receivers exist, which means that there are two maxima of (45). Therefore $f(z)=0$ has at least two roots in $\mathcal{R}_3$. Lemma 2 implies that $f(z)$ should have three zeros. Since $f(z)$ and $f'(z)$ are continuous in $\mathcal{R}_3$, $f'(z)$ must have two different zeros $\overline{z}_c$ and $\overline{z}_d$. Hence, $\Delta_-\,>0$, and (20) holds. With $\Delta_-\,>0$, we have $f(\overline{z}_c)>0$. That $f(z)$ has three zeros leads to $f(\overline{z}_d)<0$, i.e., (21) is satisfied.

Proof of T2′): While proving Lemma 2, we note that in c) of Case 3, if $\Delta_u>0$, then $z_2>2$, which means if $z_2\leq 2$, then $\Delta_u\leq 0$. When $\Delta_u\leq 0$, there is only one zero with $f'(z_b)>0$, i.e., only one CM/SW receiver exists.

Proof of T3): When the local CM/SW receiver z_d exists, we have $z_1 < \overline{z}_c < z_p < \overline{z}_d < z_q < z_2$ and $f(z_q) < 0$. Hence $z_d \in (z_q, z_2)$. Further, considering the relationship between the derivatives of l_2- and l_4-norms (Fig. 10), we have

$$-\frac{1}{z_2^3} = \frac{z_1(z_2 - z_U^{(2)})}{z_1 - z_U^{(2)}} \tag{67}$$

where $z_U^{(2)}$ represents the upper bound for z_d. The lower bound can be obtained by using the same technique. Therefore, we have

$$z_L^{(2)} = \frac{z_1(1 + z_2 z_q^3)}{1 + z_1 z_q^3}, \qquad z_U^{(2)} = \frac{z_1(1 + z_2^4)}{1 + z_1 z_2^3} \tag{68}$$

which is (25) with $\gamma_1 = 1/z_1$ and $\gamma_2 = z_2$. In order to get a tighter upper bound $z_{U'}^{(2)}$, we can substitute z_2 with $z_U^{(2)}$ and $z_U^{(2)}$ with $z_{U'}^{(2)}$ in (67). This process can be continued until the bound approximates the exact location of z_d. Due to the existence of a third-order polynomial in (67), the bound will quickly converge to z_d.

As for the global CM/SW optimum z_b, if $\overline{z}_c < 1$, $z_b \in (z_1, \overline{z}_c)$; if $\overline{z}_c \geq 1$ or $\overline{z}_c$ does not exist, $z_b \in (z_1, 1]$. Therefore, the upper bound for z_b can be $z_e = \min\{\overline{z}_c, 1\}$. Further, the two bounds can be calculated as

$$z_L^{(1)} = \frac{z_1(1 + z_1^3 z_2)}{1 + z_1^4}, \qquad z_U^{(1)} = \frac{z_1(1 + z_2 z_e^3)}{1 + z_1 z_e^3} \tag{69}$$

which gives (24).

Proof of T4): This is the direct consequence of (23). □□□

ACKNOWLEDGMENT

The authors gratefully acknowledge discussions with Professor K. Pattipati of the University of Connecticut and W. Chung of Cornell University. The authors also wish to thank the reviewers for their valuable comments and careful reading.

REFERENCES

[1] D. P. Bertsekas, *Nonlinear Programming.* Belmont, MA: Athena Scientific, 1995.
[2] W. Chung and J. LeBlanc, "The local minima of fractionally-spaced CMA blind equalizer cost function in the presence of channel noise," in *Proc. ICASSP Conf.,* Seatle, WA, May 1998, vol. VI, pp. 3345–3348.
[3] C. R. Johnson, Jr. *et. al.*, "Blind equalization using the constant modulus criterion: A review," *Proc. IEEE,* vol. 86, pp. 1927–1950, Oct. 1998.
[4] I. Fijalkow, A. Touzni, and J. R. Treichler, "Fractionally-spaced equaliztion using CMA: Robustness to channel noise and lack of disparity," *IEEE Trans. Signal Processing,* vol. 45, pp. 56–66, Jan. 1997.
[5] G. J. Foschini, "Equalizing without altering or detecting data," *Bell Syst. Tech. J.,* vol. 64, pp. 1885–1911, Oct. 1985.
[6] 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.
[7] M. Gu and L. Tong, "Geometrical characterizations of constant modulus receivers," in *Proc. 30th Conf. Inform. Sci. Syst.,* Princeton, NJ, Mar. 1998.
[8] M. Gu and L. Tong, "Power-constrained constant modulus algorithm for CDMA," in *Proc. 9th IEEE Signal Process. Workshop Stat. Array Signal Process.,* Portland, OR, Sept. 1998.
[9] 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 Process.,* pp. 301–324, 1995.
[10] 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.
[11] 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.
[12] D. Liu and L. Tong, "An analysis of constant modulus algorithm for array signal processing," *Signal Process.,* vol. 73, pp. 81–104, Jan. 1999.
[13] D. G. Luenberger, *Linear and Nonlinear Programming.* Reading, MA: Addison-Wesley, 1984.
[14] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Inform. Theory,* vol. 36, pp. 312–320, Mar. 1990.
[15] L. Tong and S.Y. Kung, "Independent component analysis in noise," in *Proc. 32nd Asilomar Conf. Signals, Syst., Comput.,* Asilomar, CA, Nov. 1998.
[16] 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.
[17] J. R. Treichler, I. Fijalkow, and C. R. Johnson, Jr., "Fractionally spaced equalizers," *IEEE Signal Processing Mag.,* pp. 45–81, May 1996.
[18] J. R. Treichler, M. G. Larimore, and J. C. Harp, "Transient and convergent behavior of the adaptive line enhancer," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-27, pp. 53–62, Feb. 1979.
[19] S. Vembu, S. Verdu, and R. A. Kennedy, "Convex cost functions in blind equalization," *IEEE Trans. Signal Processing,* vol. 42, pp. 1952–1959, Aug. 1994.
[20] H. Zeng, L. Tong, and C. R. Johnson, "Relationships between CMA and Wiener receivers," *IEEE Trans. Inform. Theory,* vol. 44, pp. 1523–1538, July 1998.
[21] S. Zeng, H. Zeng, and L. Tong, "Blind equalization using CMA: Performance analysis and a new algorithm," in *Proc. 1996 IEEE Int. Conf. Commun.,* Dallas, TX, June 1996, vol. 2, pp. 847–851.

Ming Gu (S'98) received the B.E. degree in electrical engineering in 1990 and the M.S. degree in communication and electronic systems in 1993, both from Xi'an Jiaotong University, Xi'an, China. Currently, she is pursuing the Ph.D. degree in electrical engineering at the University of Connecticut, Storrs.

Her research interests include statistical signal processing and estimation theory.

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 then was with the University of Connecticut, Storrs. Since the fall of 1998, he has been with the School of Electrical Engineering, Cornell University, Ithaca, NY, where he is now 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, adaptive receiver design for communication systems, signal processing for communication networks, and systems 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.