

# An analysis of constant modulus algorithm for array signal processing<sup>1</sup>

Dan Liu<sup>a</sup>, Lang Tong<sup>b,\*</sup>

<sup>a</sup> *Globespan Semiconductor Inc., 100 Schulz Drive, Red Bank, NJ 07701, USA*

<sup>b</sup> *School of Electrical Engineering, Cornell University, Ithaca, NY 14853, USA*

Received 25 September 1997; received in revised form 4 August 1998

---

## Abstract

The constant modulus (CM) cost function is analyzed for array signal processing. The analysis includes arbitrary source types. It is shown that CM receivers have the signal space property except in two special cases. Local minima of CM cost function are completely characterized for the noiseless case. In the presence of measurement noise, the existence of CM local minima in the neighborhood of Wiener receivers is established, and their mean squared error (MSE) performance, output power, estimation bias and residual interference are evaluated. Numerical examples are presented to demonstrate the effects of signal statistics on the locations of CM local minima. © 1999 Published by Elsevier Science B.V. All rights reserved.

## Zusammenfassung

Für die Arraysignalverarbeitung wird die Kostenfunktion des konstanten Betrages (CM) untersucht. Die Analyse schließt beliebige Quelltypen ein. Es wird gezeigt, daß CM-Empfänger mit Ausnahme von zwei Spezialfällen die Signalraumeigenschaft besitzen. Lokale Minima der CM-Kostenfunktion werden im rauschfreien Fall vollständig charakterisiert. Liegt Meßrauschen vor, dann wird die Existenz der lokalen Minima in der Nähe von Wiener-Empfängern gezeigt, und ihr mittlerer quadratischer Fehler (MSE), die Ausgangsleistung, der Schätzoffset und die Reststörung werden berechnet. Numerische Beispiele werden präsentiert, um die Effekte der Signalstatistik auf die Lage der lokalen Minima bei CM zu zeigen. © 1999 Published by Elsevier Science B.V. All rights reserved.

## Résumé

On analyse la fonction coût Module Constant (CM) pour le traitement d'antenne. L'analyse prend en compte des sources de type arbitraire. On montre que les récepteurs CM jouissent de la propriété dite de l'espace signal sauf dans deux cas particuliers. Les minima locaux de cette fonction coût CM sont complètement caractérisés en l'absence de bruit. En présence de bruit de mesure, l'existence de minima locaux au voisinage des récepteurs de Wiener est établie, et leurs performances en termes d'Erreur Quadratique Moyenne (MSE), de puissance de sortie, de biais, et d'interférence

---

\* Corresponding author. Tel.: +1 607 255 4085; fax: +1 607 255 9072; e-mail: ltong@anise.ee.cornell.edu

<sup>1</sup> This work was supported in part by the National Science Foundation under Contract NCR-9321813 and by the Office of Naval Research under Contract N00014-96-1-0895 and by the Advanced Research Projects Agency monitored by the Federal Bureau of Investigation under Contract No. J-FBI-94-221.

résiduelle sont évaluées. Des exemples numériques sont présentés pour illustrer les effets des statistiques des signaux sur les positions des minima locaux CM. © 1999 Published by Elsevier Science B.V. All rights reserved.

*Keywords:* Array processing; Beamforming; Source separation; Blind equalization and estimation

## Notations

$(\cdot)^T$	transpose
$(\cdot)^H$	Hermitian
$(\cdot)^\dagger$	Moore–Penrose inverse [8, p. 434]
$E\{\cdot\}$	expectation operator
$\ \mathbf{x}\ _p$	$p$ -norm defined by $\sqrt[p]{\sum  x_i ^p}$ .
$\ \mathbf{x}\ _A$	2-norm defined by $\sqrt{\mathbf{x}^H \mathbf{A} \mathbf{x}}$
$\mathbf{I}_n$	$n \times n$ identity matrix
$\mathbf{e}_v$	a unit column vector with 1 at the $v$ th entry and zero elsewhere
$\mathcal{C}^n$	$n$ -dimensional complex vector space
$\mathcal{C}^{n \times m}$	the set of all $n \times m$ complex matrices
$\mathcal{C}_A$	range of $\mathbf{A}\mathbf{A}^\dagger$ [8, p. 430]
$\mathcal{C}_{A^\perp}$	range of $\mathbf{I} - \mathbf{A}\mathbf{A}^\dagger$

## 1. Introduction

Since the inception of the constant modulus algorithm (CMA) in 1980, proposed first by Godard [3] and later independently by Treichler and Agee [11], CMA has been successfully implemented in various applications including equalization for microwave radio links and blind beamforming in array signal processing. Although originally invented for the equalization of a single-user intersymbol interference (ISI) channel, it has been recognized, first by Gooch and Lundell [4], that CMA applies also to the multiuser beamforming problem where the objective is to estimate one or all the source signals from an array of receivers. The so-called constant modulus array [4,5,8] is perhaps the first application of CMA in array processing.

Despite the similarity between the equalization and beam-forming problems, there are several key differences. In beam forming, different users may have different signal constellations whereas in (single user) equalization, all symbols are identically distributed. It is also possible that some interference may come from impulsive noise sources

that have super-Gaussian<sup>2</sup> statistics. Effects of signal or interference statistics on the performance of CM receivers have so far not been characterized. The second important difference is that the channel response matrix for the beam-forming problem does not have the (block) Toeplitz structure. Initialization techniques effective in the equalization problem [10] cannot be applied, which makes the initialization of CMA in beam-forming applications much more critical.

There are two classes of approaches when CMA is applied to beam-forming problems. The multi-stage CM array [6,8,9] estimates source signals sequentially. At the first stage, CMA is used to estimate one of the source signals. Which source is extracted first depends on the initialization of CMA, and it is usually unknown. The contribution of the estimated source is then subtracted from the received signal according to the mean squared error (MSE) criterion. This process continues until all sources are extracted. If the estimate at every stage is close to the minimum mean squared error (MMSE) estimate, this process approximates the MMSE estimate of all source signals obtained by the Wiener receiver; see [8,9] for the analysis. When CMA at the first stage fails to converge or its estimate has a large MSE, its error accumulates through different stages, and the multistage CMA approach may not perform well. The second class of approaches obtain the signal estimate in a single step. One example is the algebraic CMA (ACMA) proposed by Van der Veen [12]. Unlike multistage CMA, these approaches are not subject to estimation error propagation. Unfortunately, there is a lack of analysis for these algorithms about how well they behave in the presence of noise.

<sup>2</sup>A source is super- (sub-) Gaussian when  $E(|s|^4)/(E(|s|^2))^2 > (<) E(|s_G|^4)/(E(|s_G|^2))^2$  where  $s_G$  has the Gaussian distribution.

### 1.1. Main results and related work

In this paper, we perform an analysis of CMA at the first stage for the multistage CMA array approach. Such an analysis is especially important since all subsequent stages depend on the application of CMA at the first stage. Furthermore, the same analysis can also be applied to other stages when, as we shall show in this paper, the CM receiver at the first stage closely approximate the Wiener receiver. We summarize next the main results of this paper with comments of related work.

*The signal space property:* Section 3 gives the CM cost function in the general form followed by the proof of the signal space property of CM receivers. A receiver is said to have the signal space property if its coefficient vector is in the column space of the array response matrix, which is indeed the case for Wiener receivers. Such a property is significant when there are more sensors than sources. We show in Section 3 that CM receivers do have the signal space property except for in two special cases (i) there is no noise; and (ii) there are no sub-Gaussian sources. When there is no noise, there are infinitely many local minima, each can be generated from a minimum in the signal subspace. Specifically, CM local minima are made of cosets of the noise subspace generated from local minima in the signal space. Hence, it is sufficient to analyze CM local minima in the signal subspace. When there are no sub-Gaussian sources, we show that CM receivers are in the orthogonal complement of the signal subspace spanned by the columns of the array response matrix associated with the super-Gaussian sources. In other words, a CM receiver filters out all super-Gaussian sources. The signal space property of CMA is first obtained for equalization problem in [13–15]. The results presented in Section 3 are more general and its proof is more direct.

*CM local minima: the noiseless case:* The results presented in Section 4 address the following question: where are CM local minima when there is no noise? Foschini [2] was the first to show that all CM local minima inverse the channel matrix for the equalization problem. We consider here the general source condition including signals with arbitrary (sub-Gaussian, Gaussian and super-Gaussian) sources. We show that when there exist sub-Gaus-

sian sources, a receiver is a local minimum of the CM cost function if and only if it perfectly extracts a sub-Gaussian source. When sources are super-Gaussian or Gaussian, two special cases are addressed separately assuming the number of sources equals to the number of sensors. (i) If all sources are super-Gaussian, the minimum of CM cost function is unique up to a phase rotation. The output of the CM receiver in this case includes signals from all sources. (ii) If the sources are Gaussian and super-Gaussian, all receivers with a fixed output power and orthogonal to the columns of the array response matrix corresponding to the super-Gaussian sources are CM local minima. The characterization of CM local minima for the noiseless case is therefore complete.

*CM local minima: the noisy case:* In Section 5, we analyze CM local minima assuming the presence of noise in the same spirit of Zeng et al. [14]. Focusing on the case of sub-Gaussian sources, our results generalize that in [14] to include signals with different distributions. We derive tests for the existence of CM local minima in the neighborhood of Wiener receivers. These local minima are referred to as the Wiener type CM local minima. Also presented are MSE bounds, output power conditions, and properties of receiver bias and residue interference.

*Numerical examples:* The goal of Section 6 is twofold. First, we evaluate, through the use of an example, the quality of the MSE bounds derived in Section 5. Second, we study the effects of the statistics of interfering signals. We demonstrate that a super-Gaussian interference may remove a CM receiver of a sub-Gaussian source from the neighborhood of its Wiener receiver, a case that has not been reported elsewhere. Since such an example does not exist when there is no noise as shown in Section 4, it highlights the significant difference between the noiseless and noisy cases. We also demonstrate the behavior of CM receiver when sources are closely located.

## 2. The problem

Consider a multiple-input multiple-output model shown in Fig. 1:

$$\mathbf{x} = \mathbf{H}\mathbf{s} + \mathbf{w}, \quad (1)$$

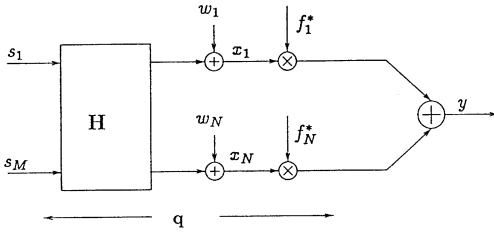


Fig. 1. The model.

where  $\mathbf{s} \triangleq (s_1, s_2, \dots, s_M)^T \in \mathcal{C}^M$  is the vector of source signals,  $\mathbf{H} \in \mathcal{C}^{N \times M}$  is the (unknown) array response matrix,  $\mathbf{w} \triangleq (w_1, w_2, \dots, w_N)^T \in \mathcal{C}^N$  is the additive noise,  $\mathbf{x} \triangleq (x_1, x_2, \dots, x_N)^T \in \mathcal{C}^N$  is the received signal vector.

A linear estimate  $y$  of one element of  $\mathbf{s}$ , say  $s_1$ , is obtained by

$$y = \mathbf{f}^H \mathbf{x} = \mathbf{q}^H \mathbf{s} + \mathbf{f}^H \mathbf{w}, \quad (2)$$

where  $\mathbf{f} \in \mathcal{C}^N$  is the estimator response vector, and  $\mathbf{q} \triangleq \mathbf{H}^H \mathbf{f} \in \mathcal{C}^M$  is the (conjugate) total system response. This model is illustrated in Fig. 1. A constant modulus (CM) estimator is defined by a local minimum of the CM cost

$$\bar{J}(\mathbf{f}) \triangleq E\{(|y|^2 - r)^2\} = E\{(|\mathbf{f}^H \mathbf{x}|^2 - r)^2\}, \quad (3)$$

where  $r$  is referred to as the dispersion constant. In general,  $r$  can be chosen arbitrarily, resulting merely a scaling factor in the local minima of the above optimization. Conventionally, if  $s_i$  is to be estimated, we set  $r = E|s_i|^4 / E|s_i|^2$  so that when there is no noise and under the general condition specified in Theorem 2,  $s_1$  can be estimated perfectly by a global minimum of Eq. (3).

In analyzing local minima of the CM cost function, we shall make the following assumptions:

A1:  $\mathbf{H} \in \mathcal{C}^{N \times M}$  has full column rank.

A2: Source vector  $\mathbf{s}$  is non-Gaussian with independent components, zero-mean with  $E(\mathbf{s}\mathbf{s}^H) = \mathbf{I}$ , and symmetrical.

A3:  $\mathbf{w}$  is Gaussian with zero mean and covariance  $\sigma^2 \mathbf{I}$ , and it is independent of  $\mathbf{s}$ .

Assumption A1 implies that the number of sensors is no less than the number of sources, and no two sources have exactly the same propagation path. The extension to non-full column rank cases can be obtained by following [15]. A2 includes the possi-

bility of sources having different statistics except the case when all sources are Gaussian. When  $\mathbf{s}$  is Gaussian, it is shown in Section 3 that, all receivers having output power equal to  $r/2$  are global minima.

### 3. The CM cost function and the signal space property

Under A2 and A3, the CM cost function in Eq. (3) has the following form:

$$\bar{J}(\mathbf{f}) = 2\|\mathbf{f}\|_R^4 - 2r\|\mathbf{f}\|_R^2 + \sum_{m=1}^M \pi_m |e_m^H \mathbf{H}^H \mathbf{f}|^4 + r^2, \quad (4)$$

where

$$\mathbf{R} \triangleq E(\mathbf{x}\mathbf{x}^H) = \mathbf{H}\mathbf{H}^H + \sigma^2 \mathbf{I}_N, \quad (5)$$

$$\pi_m \triangleq \text{cum}(s_m, s_m^*, s_m, s_m^*) = E|s_m|^4 - 2 = r_m - 2. \quad (6)$$

Some insights can be gained by examining terms in Eq. (4). Note that  $\|\mathbf{f}\|_R^2$  is the output power of the receiver. Hence, the CM cost is affected primarily by the output power of the receiver and the fourth-order terms of the system response  $e_m^H \mathbf{H}^H \mathbf{f}$  weighted by the signal cumulant  $\pi_m$ . It is evident that for a super-Gaussian source  $s_m$ ,  $\pi_m > 0$ , a CM receiver tends to minimize the corresponding term  $e_m^H \mathbf{H}^H \mathbf{f}$ . On the other hand, for a sub-Gaussian source, a CM receiver tends to enhance the corresponding term in the system response. These, of course, must be accomplished with considerations of the output power term. Note here that if all sources are Gaussian,  $\pi_m = 0$  for all  $m$ . The minimization of the CM cost leads to receivers whose output power is  $r/2$ . In general, CM estimators for this case are not useful, which is the reason we have assumed that the source vector  $\mathbf{s}$  is a non-Gaussian.

Analyzing the cost function in Eq. (4) directly is cumbersome. Most existing analyses transform the optimization of receiver coefficients to the optimization of the system response  $\mathbf{q} \triangleq \mathbf{H}^H \mathbf{f}$ . Unfortunately, the optimization of  $\mathbf{f} \in \mathcal{C}^N$  is in general *not* equivalent to the optimization of  $\mathbf{q} \in \mathcal{Q} \triangleq \{\mathbf{q} \triangleq \mathbf{H}^H \mathbf{f} | \mathbf{f} \in \mathcal{C}^N\}$  except for the case when  $\mathbf{H}$  is square and full rank. For example, when there are more sensors than sources, the analysis of CM

cost function is complicated by the existence of the null (noise) space of  $\mathbf{H}$ . The following theorem establishing the signal space property of CM receivers makes the analysis of CM receivers in  $\mathcal{Q}$  possible.

**Theorem 1.** Assume that  $\mathcal{C}_{\mathbf{H}^\perp}$  is not empty, i.e.,  $N > M$ , and  $\sigma^2 > 0$ .

1. If one of the sources is sub-Gaussian, then all local minima of  $\bar{J}(\mathbf{f})$  (4) are in the signal subspace  $\mathcal{C}_{\mathbf{H}}$  each with power satisfying

$$\|\mathbf{f}\|_{\mathbf{R}}^2 > \frac{r}{2}. \quad (7)$$

2. If one of the sources is sub-Gaussian, then all local minima of the optimization constrained by  $\mathcal{C}_{\mathbf{H}}$

$$\min_{\mathbf{f} \in \mathcal{C}_{\mathbf{H}}} \bar{J}(\mathbf{f}) \quad (8)$$

whose output power are greater than  $r/2$  are local minima of the unconstrained minimization of  $\bar{J}(\mathbf{f})$ .

3. If there is no sub-Gaussian sources, then

$$\left\{ \mathbf{f} \mid \|\mathbf{f}\|_{\mathbf{R}}^2 = \frac{r}{2}, \mathbf{f} \in \mathcal{C}_{\mathbf{H}^\perp} \cup \mathcal{C}_{\mathbf{H}_i} \right\} \quad (9)$$

is the set of global minima of  $\bar{J}(\mathbf{f})$ , where  $\mathcal{C}_{\mathbf{H}_i}$  is the space spanned by the columns of  $\mathbf{H}$  corresponding to the Gaussian sources.

**Proof.** See Appendix A.

**Remark.** CM receivers have a clear preference to signals with smaller cumulants. When there are sub-Gaussian sources, CM receivers retrieve only the sub-Gaussian sources. When there are only Gaussian and super-Gaussian sources, CM receivers filter out all super-Gaussian sources. CM receivers admit super-Gaussian signals only when all sources are super-Gaussian and  $M = N$  or  $\sigma^2 = 0$ , as shown in Theorem 2.

With Theorem 1, when there is at least one sub-Gaussian source, the optimization of  $\bar{J}(\mathbf{f})$  is equivalent to the constrained optimization in Eq. (8), which is equivalent to the optimization of  $\mathbf{q} \triangleq \mathbf{H}^H \mathbf{f} \in \mathcal{C}^M$ ,

$$J(\mathbf{q}) = 2\|\mathbf{q}\|_{\Phi}^4 - 2r\|\mathbf{q}\|_{\Phi}^2 + \sum_{m=1}^M \pi_m |q_m|^4 + r^2, \quad (10)$$

where

$$\begin{aligned} \Phi &\triangleq E\{(s + \mathbf{H}^T \mathbf{w})(s + \mathbf{H}^T \mathbf{w})^H\} \\ &= \mathbf{I}_M + \sigma^2(\mathbf{H}^H \mathbf{H})^{-1} = \mathbf{H}^T \mathbf{R}(\mathbf{H}^H)^\dagger. \end{aligned} \quad (11)$$

Eq. (10) is obtained by substituting  $\mathbf{f} = (\mathbf{H}^H)^\dagger \mathbf{q}$  into Eq. (4). If  $\mathbf{q}_c$  is a minimum of  $J(\mathbf{q})$  defined in Eq. (10), then the corresponding CM minimum can be obtained by

$$\mathbf{f}_c = (\mathbf{H}^H)^\dagger \mathbf{q}_c \quad (12)$$

whose output power is greater than  $r/2$ . Note that in the noiseless case, the set of CM local minima is the coset of  $\mathcal{C}_{\mathbf{H}^\perp}$  generated from the optimization of  $J(\mathbf{q})$ . Consequently, we shall from now on focus on the minimization of  $J(\mathbf{q})$  in  $\mathcal{C}^M$ .

#### 4. Minima of CM cost function: the noiseless case

The goal of this section is to establish the relation between the set of local minima of  $J(\mathbf{q})$  and the set of desirable system responses that completely eliminate the interference from other users. Without loss of generality, we assume that the first  $M_-$  users are sub-Gaussian followed by  $M_0$  Gaussian sources, and the last  $M_+$  signals are super-Gaussian. The complete characterization of CM local minima is given by the following theorem which generalizes that of Foschini [2].

**Theorem 2.** Consider the CM cost function  $J(\mathbf{q})$  in Eq. (10). Let  $\Lambda$  be the set of local minima of  $J(\mathbf{q})$ .

1. If there are sub-Gaussian sources ( $M_- > 0$ ), then  $\Lambda$  is the set of perfect receivers for the sub-Gaussian sources,

$$\Lambda = \left\{ q_m \mathbf{e}_m \mid |q_m|^2 = \frac{r}{r_m}, 1 \leq m \leq M_- \right\}. \quad (13)$$

2. If all sources are super-Gaussian ( $M_- = M_0 = 0$ ), then

$$\Lambda = \left\{ \mathbf{q} \mid |q_m|^2 = \frac{r}{\pi_m(2\eta_M + 1)}, m = 1, 2, \dots, M \right\}, \quad (14)$$

where

$$\eta_M \triangleq \sum_{m=1}^M \frac{1}{\pi_m}, \quad (15)$$

3. If sources are a mix of Gaussian and super-Gaussian signals ( $M_- = 0, M_0 > 0$ ), then

$$\mathcal{A} = \left\{ \mathbf{q} \mid \|\mathbf{q}\|^2 = \frac{r}{2}, q_m = 0, \forall m > M_0 \right\}. \quad (16)$$

**Proof.** See Appendix B.

**Remarks.** (1) Case 1 implies that any initialization of CM receiver leads to a perfect estimate of one of the sub-Gaussian sources. Furthermore, there exists a good initialization for every sub-Gaussian source. Sub-Gaussian signals can be estimated without being affected by the presence of Gaussian and super-Gaussian signals. Gaussian and super-Gaussian signals cannot be estimated by CMA when there exist sub-Gaussian sources. Fig. 2 illustrates those properties for the three-source cases when one, two or three of the sources are sub-Gaussian. Let

$$\mathbf{q}^1 = \begin{pmatrix} q_1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{q}^2 = \begin{pmatrix} 0 \\ q_2 \\ 0 \end{pmatrix}, \quad \mathbf{q}^3 = \begin{pmatrix} 0 \\ 0 \\ q_3 \end{pmatrix},$$

where  $|q_m|^2 = r/r_m, m = 1, 2, 3$ .

First, as shown in Fig. 2(a), if only  $s_1$  is sub-Gaussian, the CM cost function has a local minimum  $\mathbf{q}^1$ . If there are two sub-Gaussian sources  $s_1, s_2$ , CM receivers are  $\mathbf{q}^1$  and  $\mathbf{q}^2$ , shown in Fig. 2(b). Finally, the CM receivers are  $\mathbf{q}^1, \mathbf{q}^2$  and  $\mathbf{q}^3$ , when all the three sources are sub-Gaussian (Fig. 2(c)).

(2) When  $N = M$  and all sources are super-Gaussian (Case 2), local minima are located at a vertices of a box as illustrated in Fig. 3(a).

(3) For Case 3, in which there is no sub-Gaussian sources, CM receivers are on a quadratic surface in the subspace spanned by the columns of  $\mathbf{H}$  associated with Gaussian sources. For example, if the first two sources are Gaussian and the third is super-Gaussian, any

$$\mathbf{q} = \begin{pmatrix} q_1 \\ q_2 \\ 0 \end{pmatrix}$$

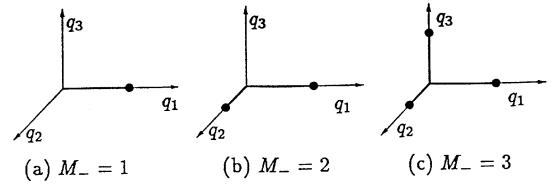


Fig. 2. CM receivers for case 1 ( $M = 3$ ).

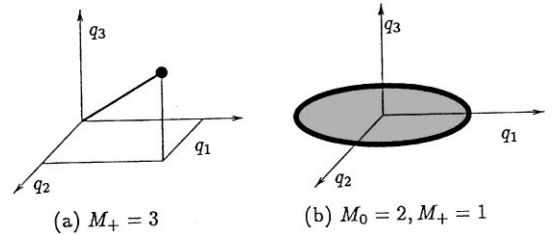


Fig. 3. CM Receivers for cases 2 and 3 ( $M = 3$ ).

with  $q_1^2 + q_2^2 = r/2$  are the local minima of CM cost function (51), which is illustrated in Fig. 3(b).

## 5. Minima of CM cost function: the noisy case

The presence of noise complicates the analysis considerably. In such a case, the characterization of all local minima of CM cost in ways similar to Theorem 2 appears to be difficult. Our goals in this section are (i) to locate those local minima near the Wiener receivers; and (ii) to investigate the properties of these local minima including their output power, bias and MSE. Following [13,14] where it was shown that, when the MMSE is not too large, there exists a CM local minimum in the small neighborhood of the Wiener receiver, we obtain similar results for complex signals with possibly different distributions. To simplify the analysis, we shall further restrict ourselves to the cases when all sources are sub-Gaussian, i.e.

$$A4: E(|s_m|^4) < 2, \quad m = 1, 2, \dots, M.$$

### 5.1. Location of CM receivers

To locate CM local minima in the neighborhood of Wiener receivers, we adopt the two-step

approach by Zeng and Tong in [13]. First, a compact neighborhood of the Wiener receiver is defined using parameters relevant to the receiver MSE performance. Second, parameters are chosen such that the CM cost function on the boundary of the neighborhood is strictly greater than a reference inside the neighborhood. By the Weierstrass theorem, this neighborhood must contain at least one local minimum. Before we proceed in defining the neighborhood that captures a CM receiver, let us recall the Wiener receiver defined as one that minimizes the mean square error

$$J_w(\mathbf{q}) \triangleq E(|y - s_1|^2) = \|\mathbf{q}\|_{\Phi}^2 - \mathbf{q}^H \mathbf{e}_1 - \mathbf{e}_1^H \mathbf{q} + 1. \quad (17)$$

The Wiener receiver for signal  $s_1$  is the minimum of the above optimization

$$\mathbf{q}_w = \arg \min J_w(\mathbf{q}) = \Phi^{-1} \mathbf{e}_1. \quad (18)$$

Again, without loss of generality, we shall restrict our discussion of estimating  $s_1$  throughout this section.

### 5.1.1. The neighborhood

The neighborhood is defined through a special parameterization of the system response  $\mathbf{q}$ . Given an estimator of  $s_1$ , denote its system response by

$$\mathbf{q} \triangleq \theta \begin{pmatrix} 1 \\ \mathbf{q}_1 \end{pmatrix}, \quad (19)$$

where  $\theta$  is the receiver gain of  $s_1$  and  $\mathbf{q}_1$  is the vector of coefficients corresponding to the (relative) residue interference from other users. Suppose the Wiener receiver

$$\mathbf{q}_w = \theta_w \begin{pmatrix} 1 \\ \mathbf{q}_{w1} \end{pmatrix}. \quad (20)$$

The neighborhood  $\mathcal{B}(\mathbf{q}_w, \delta_U, \theta_L, \theta_U)$ , illustrated in Fig. 4, is a slice of a cone centered around the Wiener receiver  $\mathbf{q}_w$  with receiver gain  $\theta \in [\theta_L, \theta_U]$ ,

$$\mathcal{B}(\mathbf{q}_w, \delta_U, \theta_L, \theta_U) \triangleq \left\{ \mathbf{q} \triangleq \theta \begin{pmatrix} 1 \\ \mathbf{q}_1 \end{pmatrix} : \theta_L \leq \theta \leq \theta_U, \right. \\ \left. \|\mathbf{q}_1 - \mathbf{q}_{w1}\|_C \triangleq \delta \leq \delta_U \right\}, \quad (21)$$

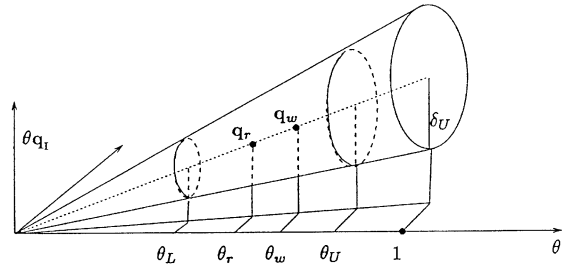


Fig. 4. The region  $\mathcal{B}$ .

where  $C$  is the submatrix of  $\Phi$  by removing the first column and row. The ‘width’ of the cone is specified by  $\delta_U$  which is, as shown in [14], the maximum extra unbiased MSE<sup>3</sup> among all receivers in the neighborhood, i.e.,

$$\delta_U^2 = \max_{\mathbf{q} \in \mathcal{B}(\mathbf{q}_w, \delta_U, \theta_L, \theta_U)} \left\{ \text{MSE} \left( \frac{1}{\theta_w} \mathbf{q}_w \right) - \text{MSE} \left( \frac{1}{\theta} \mathbf{q} \right) \right\}. \quad (22)$$

Here  $\text{MSE}(\mathbf{q})$  stands for the mean square error of an estimator with system response  $\mathbf{q}$ .

### 5.1.2. The location of CM local minima

Select the reference  $\mathbf{q}_r = \theta_r \begin{pmatrix} 1 \\ \mathbf{q}_{r1} \end{pmatrix}$  where  $\theta_r$  is such that  $J \left( \theta_r \begin{pmatrix} 1 \\ \mathbf{q}_{r1} \end{pmatrix} \right)$  is minimized, i.e.,

$$\theta_r = \arg \min_{\theta} J \left( \theta \begin{pmatrix} 1 \\ \mathbf{q}_{r1} \end{pmatrix} \right), \quad (23)$$

$$\mathbf{q}_r = \theta_r \begin{pmatrix} 1 \\ \mathbf{q}_{r1} \end{pmatrix}. \quad (24)$$

We aim to choose  $\theta_L$ ,  $\theta_U$  and  $\delta_U$  so that (i)  $\mathbf{q}_r$  is inside the neighborhood; and (ii) the CM cost on the boundary of  $\mathcal{B}(\mathbf{q}_w, \delta_U, \theta_L, \theta_U)$  is strictly greater

<sup>3</sup>Given an estimator of  $s_1$ , let the system response be  $\mathbf{q} = \theta \begin{pmatrix} 1 \\ \mathbf{q}_1 \end{pmatrix}$ . The estimate  $y$  of  $s_1$  has the form  $y = \mathbf{q}^T \mathbf{s} + \text{noise term} = \theta s_1 + \text{other terms orthogonal to } s_1$ . Because  $E(y|s_1) = \theta s_1$ ,  $y$  is a conditionally biased estimate of  $s_1$ . Scaling the estimate  $y$  by  $1/\theta$ , we have a (conditionally) unbiased estimate  $u \triangleq (1/\theta)y$  of  $s_1$ , i.e.  $E(u - s_1|s_1) = 0$ . See also [1].

than  $J(\mathbf{q}_r)$ . Consequently, a CM receiver is located in  $\mathcal{B}(\mathbf{q}_w, \delta_U, \theta_L, \theta_U)$ . The following theorem describes such a procedure.

**Theorem 3.** Given the Wiener receiver  $\mathbf{q}_w = \theta_w \begin{pmatrix} 1 \\ \mathbf{q}_{w1} \end{pmatrix}$ , let  $\mathbf{q}_r$  be defined by Eq. (24). With cumulants  $\pi_i$  of the sources, define

$$\pi_{\min} \triangleq \min_{1 \leq m \leq M} \{\pi_m\}, \quad r_{\min} \triangleq \pi_{\min} + 2, \quad (25)$$

$$\mathbf{\Pi}_1 \triangleq \text{diag} \left\{ \left( \frac{\pi_2}{\pi_{\min}} \right)^{1/4}, \dots, \left( \frac{\pi_M}{\pi_{\min}} \right)^{1/4} \right\}. \quad (26)$$

Let  $\theta_L$  and  $\theta_U$  be determined by

$$\theta_L \triangleq \min_{0 \leq \delta \leq \delta_U} \sqrt{\frac{-c_1(\delta) - c_1(\delta)^2 - 4c_2(\delta)c_0}{2c_2(\delta)}}, \quad (27)$$

$$\theta_U \triangleq \max_{0 \leq \delta \leq \delta_U} \sqrt{\frac{-c_1(\delta) + c_1(\delta)^2 - 4c_2(\delta)c_0}{2c_2(\delta)}}, \quad (28)$$

where

$$c_0 \triangleq \frac{r^2}{2 + \pi_1 \theta_w^2 + \pi_{\min} \theta_w^2 \|\mathbf{\Pi}_1 \mathbf{q}_{w1}\|_4^4} \quad (29)$$

$$c_1(\delta) \triangleq -2r \left( \delta^2 + \frac{1}{\theta_w} \right), \quad (30)$$

$$c_2(\delta) \triangleq 2 \left( \delta^2 + \frac{1}{\theta_w} \right)^2 + \pi_1 + \pi_{\min} (\delta + \|\mathbf{\Pi}_1 \mathbf{q}_{w1}\|_4)^4. \quad (31)$$

Let  $\delta_U$  be the smallest positive root of  $D(\delta) \triangleq c_1(\delta)^2 - 4c_2(\delta)c_0$ .

(a) With dispersion constant  $r = E|s_1|^4 / E|s_1|^2 = E|s_1|^4$ , if

$$\theta_w > \max \left\{ 1 - \frac{2}{3(2 - r_{\min})}, 1 - \frac{\sqrt{4 + 16r - 9rr_{\min} - r + 2}}{20 - 9r_{\min} - r} \right\} \quad (32)$$

and

$$\delta_U < \|\mathbf{q}_{w1}\|_2, \quad (33)$$

then there exists a local minimum of  $J(\mathbf{q})$  in  $\mathcal{B}(\mathbf{q}_w, \delta_U, \theta_L, \theta_U)$ .

(b) Let  $\mathbf{q}_c$  is the local minimum of  $J(\mathbf{q})$  in  $\mathcal{B}(\mathbf{q}_w, \delta_U, \theta_L, \theta_U)$ , then

$$\lim_{\sigma \rightarrow 0} \mathcal{B}(\mathbf{q}_w, \delta_U, \theta_L, \theta_U) = \lim_{\sigma \rightarrow 0} \{\mathbf{q}_c\} = \lim_{\sigma \rightarrow 0} \{\mathbf{q}_w\}. \quad (34)$$

**Proof.** See Appendix C.

Theorem 3 gives sufficient conditions of the existence of a CM receiver in the neighborhood of the Wiener receiver. Given the Wiener receiver and the distribution of source signals, one can check the two conditions given in Eqs. (32) and (33), both depend on the distributions of source signals. If they are satisfied, then there exists a CM receiver in the neighborhood of the Wiener receiver. Table 1 gives the maximum value of MSE of the Wiener receiver for QAM sources to satisfy Eq. (32) assuming that  $r = r_{\min}$ . As to Eq. (33), polynomial  $D(\delta)$  needs to be calculated first, then its smallest positive root  $\delta_U$ . When the noise is small enough, for any sources, a root of  $D(\delta)$  can always be found less than  $\|\mathbf{q}_{w1}\|_2$  and Eq. (32) can be satisfied with  $\theta_w$  close to 1, which means that there must be a local minimum of  $J(\mathbf{q})$  in the neighborhood of the Wiener receiver. Part (b) of Theorem 3 gives the result consistent with the noiseless case. As measurement noise becomes zero, the neighborhood defined by  $\mathcal{B}(\mathbf{q}_w, \delta_U, \theta_L, \theta_U)$  shrinks to one, the Wiener receiver. We note that the above theorem does not cover those local minima not in the vicinity of Wiener receivers.

## 5.2. Output power, bias and MSE of CM receiver

In Section 3, it is shown that the output power of any CM receiver is greater than  $r/2$  when there is at least one sub-Gaussian source. In this section, the upper bound of the output power of the local minima of  $J(\mathbf{q})$  is obtained for the case that all the sources are sub-Gaussian. Further, we compare the output power, bias, residue interference of CM receiver in the region  $\mathcal{B}(\mathbf{q}_w, \delta_U, \theta_L, \theta_U)$  with that of the Wiener receiver, and give the MSE bounds of the CM receiver.



Table 1  
Required MMSE for the test of existence of CM receivers in the neighborhood of Wiener receivers

	$r = E\{ s ^4\}$	Required MMSE
4QAM	1.000	0.5683
16QAM	1.320	0.4482
32QAM	1.310	0.4529
64QAM	1.384	0.4156
128QAM	1.348	0.4344

**Theorem 4.** (1) *The output power of any CM receiver  $f_c$  satisfies*

$$\frac{r}{2} < \|f_c\|_R^2 < \frac{r}{r_{\min}}. \quad (35)$$

(2) *Suppose that  $\mathbf{q}_c = \theta_c [1, \mathbf{q}_{c1}]^t \in \mathcal{B}(\mathbf{q}_w, \delta_U, \theta_L, \theta_U)$  is a CM receiver,*

(a) *If*

$$\theta_w > \max \left\{ \frac{2}{3^2}, \frac{2r}{2 + r_{\min}} \right\}, \quad (36)$$

and

$$\theta_w^3 - \theta_w^2 - \frac{r_{\min}}{4 - 2r_{\min}} \theta_w + \frac{r}{4 - 2r_{\min}} < 0, \quad (37)$$

*then the output power of CMA receiver is less than that of the Wiener receiver, i.e.*

$$\|\mathbf{q}_c\|_{\Phi}^2 < \|\mathbf{q}_w\|_{\Phi}^2 < 1, \quad (38)$$

*and the bias of the CMA receiver is greater than that of the Wiener receiver, i.e.*

$$|\theta_c| \leq \theta_w. \quad (39)$$

(b)  $\|\mathbf{q}_{c1}\|_4 \geq \|\mathbf{q}_{w1}\|_4$ .

**Proof.** See Appendix D.

With  $r_{\min} = r$ , Eq. (37) is equivalent to

$$\theta_w > \sqrt{r/4 - 2r}. \quad (40)$$

For some signals, Eq. (40) can never be true. For example, 64QAM and 128QAM,  $\sqrt{r/(4 - 2r)} = 1.0599$  and  $\sqrt{r/(4 - 2r)} = 1.0167$ , respectively. Our

simulation shows that the output power of CMA receiver is greater than that of the Wiener receiver for 64QAM and 128QAM sources, respectively. If Eq. (37) is true, we also can know that  $|\theta_r| < \theta_w$ , i.e., the bias of the reference  $\mathbf{q}_r$  is larger than that of Wiener receiver.

A useful performance measure of a receiver is its mean squared error (MSE). In the following theorem, the MSE bounds are given for the local minimum located in  $\mathcal{B}(\mathbf{q}_w, \delta_U, \theta_L, \theta_U)$ .

**Theorem 5.** *Suppose  $\mathbf{q}_w = \theta_w \begin{pmatrix} 1 \\ \mathbf{q}_{w1} \end{pmatrix}$  is the Wiener receiver,  $\mathbf{q}_c$  is a local minimum in  $\mathcal{B}(\mathbf{q}_w, \delta_U, \theta_L, \theta_U)$  without phase rotation, and  $\mathbf{q}_r = \theta_r \begin{pmatrix} 1 \\ \mathbf{q}_{r1} \end{pmatrix}$  is the reference defined in Eqs. (23) and (24). Let  $\Delta^{\mathcal{E}} \triangleq J_w(\mathbf{q}_c) - J_w(\mathbf{q}_w)$ , and  $\widehat{\Delta}^{\mathcal{E}} \triangleq J_w(\mathbf{q}_r) - J_w(\mathbf{q}_w)$ ; then*

$$\begin{aligned} & \min \underbrace{\left\{ \frac{(\theta_U - \theta_w)^2}{\theta_w}, \frac{(\theta_L - \theta_w)^2}{\theta_w} \right\}}_{\Delta^{\mathcal{E}_L}} \\ & \leq \Delta^{\mathcal{E}} \\ & \leq \underbrace{\max \left\{ \frac{(\theta_L - \theta_w)^2}{\theta_w}, \frac{(\theta_U - \theta_w)^2}{\theta_w} \right\}}_{\Delta^{\mathcal{E}_U}} + (\theta_U \delta_U)^2, \end{aligned} \quad (41)$$

$$\begin{aligned} \widehat{\Delta}^{\mathcal{E}} &= (\sqrt{r/(2 + \pi_1 \theta_w^2 + \pi_{\min} \theta_w^2 \|\mathbf{\Pi}_1 \mathbf{q}_{w1}\|_4^4)} - \sqrt{\theta_w})^2 \\ &= \left( \frac{3r - 4}{2r} \right)^2 J_w^2(\mathbf{q}_w) + O(J_w^2(\mathbf{q}_w)), \end{aligned} \quad (42)$$

where  $J_w(\mathbf{q})$  is the MSE of a receiver  $\mathbf{q}$ .

**Proof.** See Appendix E.

Similar to the conditions (32) and (33) in Theorem 3,  $\Delta^{\mathcal{E}_U}$ ,  $\Delta^{\mathcal{E}_L}$  and  $\widehat{\Delta}^{\mathcal{E}}$  can be computed from the Wiener receiver and the distributions of sources. If  $J_w(\mathbf{q}_w)$ , the MSE of Wiener receiver, is small, the extra MSE of CM receiver over that of the Wiener receiver can be approximated by Eq. (42).

## 6. Numerical results

We present in this section several numerical studies of CM cost function when it is applied to estimating a narrow-band source using a linearly spaced antenna array. We have focused our study on the case with channel noise since the CM local minima are completely characterized in Theorem 2 for the noiseless case. The objective of this section is twofold. First, we evaluate the quality of the MSE bounds derived in Section 5. Second, we investigate the effect of the presence of super-Gaussian interference. In contrast to the noiseless case, the presence of super-Gaussian interference may remove the CM local minimum from the neighborhood of Wiener receiver.

We tested the case with four uniformly quarter-wavelength spaced antenna elements and three sources, i.e.,  $N = 4$ ,  $M = 3$ . All users had the same power. We considered the performance of CM receiver for (i) well separated sources; (ii) closely spaced interference with different source distributions; and (iii) clustered sources.

### 6.1. Properties of CM local minima

We considered first the case of three well-separated sources: a QPSK source at  $-10^\circ$ , a 16-QAM source at  $15^\circ$ , and a 64-QAM at  $35^\circ$ . To estimate QPSK source  $s_1$  at  $-10^\circ$ , we chose  $r = r_1 = 1$ . Fig. 5 shows the lower and upper MSE bounds given in Eq. (41), the MSE of the Wiener receiver, and the MSE of the CM receiver.<sup>4</sup> We also marked the reference point  $\mathbf{q}_r$ , defined in Eqs. (23) and (24), as an approximation of the CM receiver. From Fig. 5, we observe that both the lower and the upper bounds were tight. The MSEs of CM receiver  $\mathbf{q}_c$  (-), its approximation  $\mathbf{q}_r$  (\*), and the Wiener receiver  $\mathbf{q}_w$  (-) were indeed very close to each other. This was further demonstrated in Table 2 for SNR = 10 dB. For this case, we see that the output power of CMA is less than that of the Wiener receiver, and the interference level is greater than

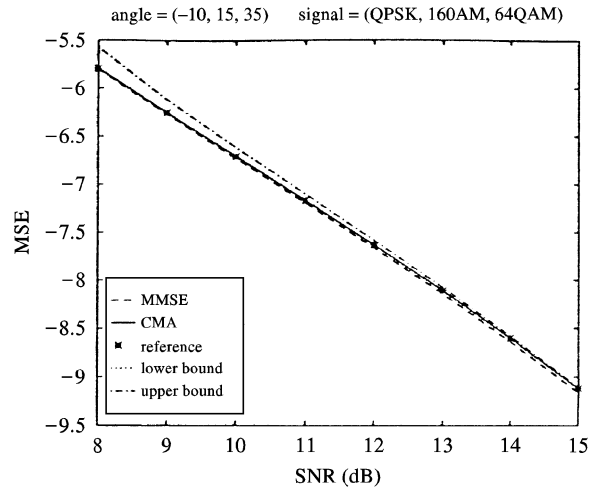


Fig. 5. The bounds of MSE.

that of the Wiener receiver, both predicted by Theorem 4. Note that because the output power is greater than 0.5, the local minimum obtained by searching  $\mathbf{q}$  is indeed the local minimum obtained by searching  $\mathbf{f}$  as shown in Theorem 1. When the condition (40) of Theorem 4 is not satisfied, the output power of the CM receiver may be greater than that of the Wiener receiver, which is in contrast to the equalization problem for BPSK signals [14]. Table 3 documents one such case when all three signals were 64-QAM sources. Note also that the CMA, its approximation, and the Wiener receiver remained close in this case.

### 6.2. Effects of co-channel interference

We considered the same three source case but with varied interference statistics, QPSK, 16QAM, 64QAM, Gaussian sources and super-Gaussian sources with  $r = E|s|^4 = 3$ . It is evident from Fig. 6 that, for the case of well-separated sources, the signal statistics of the interference affects little the location of CM local minimum and its approximations.

#### 6.2.1. Can super-Gaussian interference remove a CM local minimum?

We considered next the case when the source locations were closer. Specifically, the angles of

<sup>4</sup>The CM receiver is obtained from the gradient search of the true CM cost function.

Table 2  
CMA and MMSE at SNR = 10 dB ( $[-10^\circ, 15^\circ, 35^\circ]$ , [QPSK, 16QAM, 64QAM])

	System response coefficients	Output power	$\ q_1\ _4^4$
CMA $q_c$	[0.757, 0.122 + 0.198i, 0.021 - 0.108i]	0.7269	0.0094
MMSE $q_m$	[0.788, 0.123 + 0.200i, 0.022 - 0.118i]	0.7878	0.0083
Reference $q_r$	[0.757, 0.118 + 0.192i, 0.022 - 0.113i]	0.7268	0.0083

Table 3  
CMA and MMSE at SNR = 10 dB ( $[-10^\circ, 15^\circ, 35^\circ]$ , [64QAM, 64QAM, 64QAM])

	System response coefficients	Output power	$\ q_1\ _4^4$
CMA $q_c$	[0.822, 0.135 + 0.219i, 0.022 - 0.114i]	0.8574	0.0101
MMSE $q_m$	[0.788, 0.123 + 0.200i, 0.022 - 0.118i]	0.7878	0.0083
Reference $q_r$	[0.822, 0.128 + 0.208i, 0.024 - 0.123i]	0.8572	0.0083

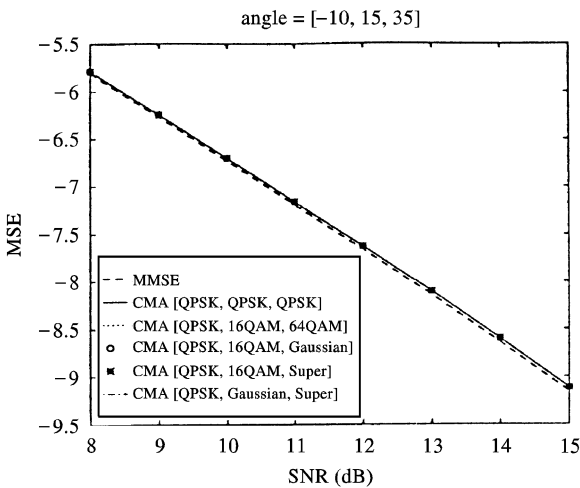


Fig. 6. The effect of interference signal constellation.

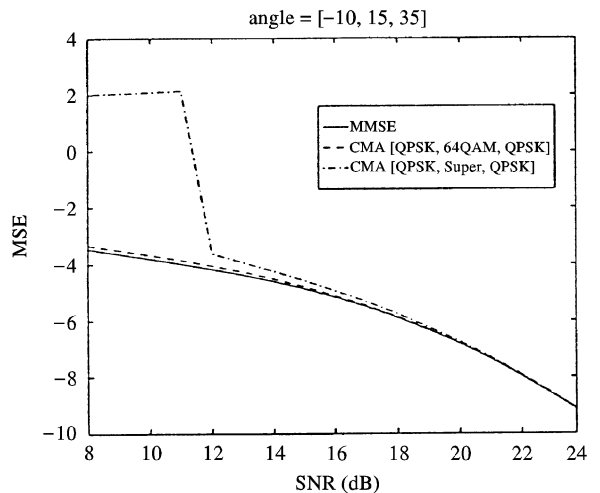


Fig. 7. The effect of interference signal constellation.

arrival were  $-10^\circ, 0^\circ$  and  $35^\circ$ , which means that interference from the second user was closer to the desired signal. Fig. 7 shows the MSE of CM receiver for the QPSK source  $s_1$  with the interference  $s_2$  varied from 64-QAM to a super-Gaussian source  $r_2 = 6$ . In this case, we observed that the effects of interference constellation is not significant when they were sub-Gaussian. However, when the interference  $s_2$  was changed to super-Gaussian with  $r_2 = 6$ , the MSE behavior of CM receiver at lower

SNR region showed significant difference from the case when  $s_2$  was 64-QAM. For the noiseless case, we have already shown in Theorem 2 that the presence of super-Gaussian interference has no effect on the location of the CM receiver for a sub-Gaussian source, which was indeed verified at high SNR region. At low SNR, this example appeared to show the contrary. Table 4 gives the CM coefficients for the three users at SNR = 10 dB. It appeared that, initialized at their Wiener receivers, all

Table 4  
CMA and MMSE (angles:  $[-10^\circ, 0^\circ, 35^\circ]$ , SNR = 10 dB)

SNR = 10 dB		
user 1	CMA	$[0.016 + 0.082i, 0.031 - 0.140i, 0.861]$
$r = 1$	MMSE	$[0.583, 0.370 + 0.161i, 0.016 - 0.082i]$
user 2	CMA	$[0.016 + 0.085i, 0.028 - 0.127i, 0.826]$
$r_2 = 6$	MMSE	$[0.370 - 0.161i, 0.486, 0.031 + 0.140i]$
user 3	CMA	$[0.016 + 0.085i, 0.028 - 0.127i, 0.826]$
$r_3 = 1$	MMSE	$[0.017 + 0.082i, 0.031 - 0.139i, 0.861]$

CM receivers converged to the one close to the Wiener receiver for user  $s_3$ . Further test<sup>5</sup> showed that it seemed to be the case that the super-Gaussian interference did remove the local minimum from the neighborhood of the Wiener receiver for user 1.

6.3. Closely spaced sources

Finally, we considered the case when the three sources were clustered with angles of arrival at 10, 15 and 20°. All source signals were QPSK. Figs. 8 and 9 show the MSE of CM receivers for user 1 and user 2, respectively. The presence of noise affects these two users quite differently. At low SNR, there was no CM local minimum in the neighborhood of Wiener receiver for user 1 due to the power constraint given in Theorem 1. Table 5 shows that, when initialized at the Wiener receiver, the gradient search leads to one close to the Wiener receiver for user 2. As SNR increases, the CM local minimum re-emerges in the neighborhood of Wiener receiver of user 1. This phenomenon does not occur for user 2 for SNR from 8 to 10 dB, and it occurs for SNR in the range of 11–43 dB. Shown in Table 6 is the case for SNR = 14 dB. Initialized at the Wiener receiver for user 2, the CM gradient search leads to the local minimum in the neighborhood of Wiener receiver for user 3.

<sup>5</sup> We checked the gradient of CM cost function in the neighborhood of the Wiener receiver. Among receivers having less than 50% extra MSE, i.e.  $(MSE(q) - MSE(q_m))/MSE(q_m) < 50\%$ , 10<sup>6</sup> receivers were tested for their gradient and found them lower bounded by 0.0005.

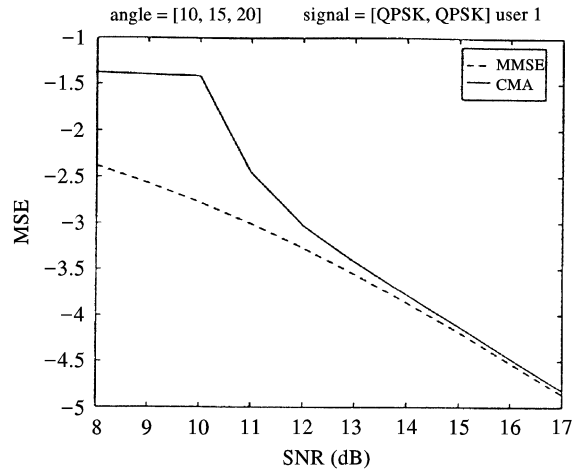


Fig. 8. The MSE of closely spaced sources (user 1).

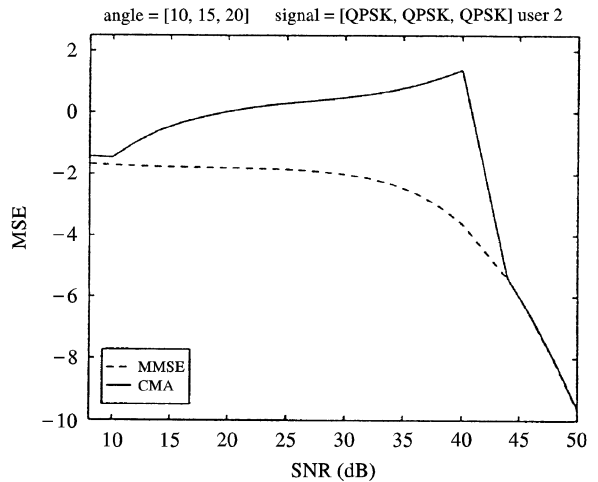


Fig. 9. The MSE of closely spaced sources (user 2).

Table 5  
CMA and MMSE (angles:  $[10^\circ, 15^\circ, 20^\circ]$ , SNR = 8 dB)

SNR = 8 dB		
user 1	CMA	$[0.421 - 0.086i, 0.435, 0.422 + 0.084i]$
$r_1 = 1$	MMSE	$[0.422, 0.310 + 0.063i, 0.191 + 0.080i]$
user 2	CMA	$[0.421 - 0.086i, 0.435, 0.422 + 0.084i]$
$r_2 = 1$	MMSE	$[0.310 + 0.063i, 0.321, 0.312 + 0.620i]$
user 3	CMA	$[0.420 - 0.085, 0.435, 0.423 + 0.084i]$
$r_3 = 1$	MMSE	$[0.191 - 0.080i, 0.312 - 0.062i, 0.420]$

Table 6  
CMA and MMSE (angles:  $[10^\circ, 15^\circ, 20^\circ]$ , SNR = 14 dB)

SNR = 14 dB		
user 1	CMA	$[0.600, 0.361 + 0.073i, 0.125 + 0.052i]$
$r_1 = 1$	MMSE	$[0.589, 0.321 + 0.065i, 0.061 + 0.026i]$
user 2	CMA	$[0.129 - 0.054i, 0.368 - 0.073i, 0.597]$
$r_2 = 1$	MMSE	$[0.321 - 0.067i, 0.335, 0.325 + 0.065i]$
user 3	CMA	$[0.127 - 0.053i, 0.367 - 0.073i, 0.598]$
$r_3 = 1$	MMSE	$[0.061 - 0.026i, 0.325 - 0.065i, 0.585]$

## 7. Conclusion

Blind beam forming based on the constant modulus cost function has several distinct advantages, especially when all sources are sub-Gaussian. The most important is, perhaps, the connection between the Wiener and CM receivers. When there is no noise and all sources are sub-Gaussian, all local minima of CM cost function are global and they perfectly estimate all the sources. When there is noise, we have shown in this paper that those sources whose MMSE is not too large can also be estimated by CM receiver with similar MSE. In fact, CM receivers for these sources are approximately scaled Wiener receivers. This means that the CM array approach is close to the MMSE estimate of all the sources provided that, at each stage, the MMSE of some source is not too large and appropriate initialization is used. There are cases, however, that there is no CM receiver in the neighborhood of the Wiener receiver. In this case, the CM array cannot extract this source accurately, and significant error propagation may occur.

When there are mixed sources, it should be cautioned that the presence of super-Gaussian sources may remove CM local minima from the neighborhood of Wiener receivers at low SNR. When there is no noise and at least one sub-Gaussian source, we show that CM receivers filter out all Gaussian and super-Gaussian sources. When all sources are either Gaussian or super-Gaussian, CM receivers filter out all super-Gaussian sources. It is perhaps not surprising that CM cost function tends to favor sub-Gaussian over Gaussian and super-Gaussian sources, and Gaussian sources over super-Gaussian sources.

For further reading, see [3].

## Acknowledgements

The authors wish to thank the anonymous reviewer for the careful reading and comments that helped to improve the presentation of the paper.

## Appendix A. Proof of Theorem 1

To prove case 1, assume, without the loss of generality, that  $\pi_1 < 0$ . Let  $\mathbf{f}^0$  be a local minima, and  $\Delta \mathbf{f} = \Delta \mathbf{f}_H + \Delta \mathbf{f}_{H^\perp}$ , then

$$\begin{aligned} \bar{J}(\mathbf{f}^0) - \bar{J}(\mathbf{f}^0 + \Delta \mathbf{f}) &= \{2(\|\mathbf{f}\|_R^4 - \|\mathbf{f} + \Delta \mathbf{f}\|_R^4) \\ &\quad - 2r(\|\mathbf{f}^0\|_R^2 - \|\mathbf{f}^0 + \Delta \mathbf{f}\|_R^2)\} \end{aligned} \quad (\text{A.1})$$

$$+ \left\{ \sum_{m=1}^M \pi_m (|e_m^H \mathbf{H}^H \mathbf{f}^0|^4 - |e_m^H \mathbf{H}^H (\mathbf{f}^0 + \Delta \mathbf{f}_H)|^4) \right\} \quad (\text{A.2})$$

$$= \Delta_1 + \Delta_2. \quad (\text{A.3})$$

To show that all CM local minima are in the signal subspace, we consider the following two scenarios: (a)  $\mathbf{f}^0 = \mathbf{f}_{H^\perp}^0 \in \mathcal{C}_{H^\perp}$ ; (b)  $\mathbf{f}^0 = \mathbf{f}_H^0 + \mathbf{f}_{H^\perp}^0, \mathbf{f}_H^0 \in \mathcal{C}_H, \mathbf{f}_{H^\perp}^0 \in \mathcal{C}_{H^\perp}, \mathbf{f}_{H^\perp}^0 \neq 0$ .

For (a), because  $\pi_1 < 0$ , and (A1), there exists  $\Delta \mathbf{f}_H \in \mathcal{C}_H$  and an arbitrarily small  $\varepsilon$  such that

$$\begin{aligned} \Delta_2 &= \sum_{m=1}^M \pi_m (|e_m^H \mathbf{H}^H \mathbf{f}_{H^\perp}^0|^4 \\ &\quad - |e_m^H \mathbf{H}^H (\mathbf{f}_{H^\perp}^0 + \Delta \mathbf{f}_H + \Delta \mathbf{f}_{H^\perp})|^4) \\ &= - \sum_{m=1}^M \pi_m |e^H \mathbf{H}^H \Delta \mathbf{f}_H|^4 = -\varepsilon \pi_1 > 0. \end{aligned} \quad (\text{A.4})$$

Since  $\mathbf{f}_{H^\perp}^0 \neq 0$ , and  $\sigma^2 > 0$ , there exists a  $\Delta \mathbf{f}_{H^\perp} \in \mathcal{C}_{H^\perp}$  such that

$$\begin{aligned} \|\mathbf{f}^0\|_R^2 - \|\mathbf{f}^0 + \Delta \mathbf{f}\|_R^2 \\ = \|\mathbf{f}_{H^\perp}^0\|_R^2 - \|\mathbf{f}_{H^\perp}^0 + \Delta \mathbf{f}_{H^\perp}\|_R^2 - \|\Delta \mathbf{f}_H\|_R^2 = 0, \end{aligned}$$

i.e.,

$$\Delta_1 = 0. \quad (\text{A.5})$$

From Eqs. (A.5) and (A.6), we have  $\bar{J}(\mathbf{f}^0) > \bar{J}(\mathbf{f}^0 + \Delta\mathbf{f})$ . Therefore,  $\mathbf{f}^0$  can not be a local minimum. The same technique can also be used to show that  $\mathbf{f}_{H^\perp}^0 = \mathbf{0}$  for case (b).

We show next that all CM local minima in this case must have output power greater than  $r/2$ . Let  $\mathbf{f}_H^0$  be a CM local minimum and  $\|\mathbf{f}_H^0\|_R^2 < r/2$ . Let  $\mathbf{f} = \Delta\mathbf{f}_{H^\perp} + \mathbf{f}_H^0$ , where  $\Delta\mathbf{f}_{H^\perp} \in \mathcal{C}_{H^\perp}$ . It can be verified that

$$\bar{J}(\mathbf{f}) = \bar{J}(\mathbf{f}_H^0) + 2\|\Delta\mathbf{f}_{H^\perp}\|_R^4 + 2\|\Delta\mathbf{f}_{H^\perp}\|_R^2(2\|\mathbf{f}_H^0\|_R^2 - r). \quad (\text{A.6})$$

Because the last term of the above equation is negative, there exists  $\Delta\mathbf{f}_{H^\perp}$  small enough that  $\bar{J}(\mathbf{f}) < \bar{J}(\mathbf{f}_H^0)$ . Hence  $\|\mathbf{f}_H^0\|_R^2 \geq r/2$ . To exclude the equality, one can always find a  $\mathbf{f}_H^0 + \Delta\mathbf{f}_H$  such that  $(2\|\mathbf{f}_H^0 + \Delta\mathbf{f}_H\|_R^2 - r) < 0$ .

To show case 2, we show that any local minimum in  $\mathcal{C}_H$  with output power greater than  $r/2$  is also a local minimum with no constraint. Consider a small perturbation  $\Delta\mathbf{f} = \Delta\mathbf{f}_H + \Delta\mathbf{f}_{H^\perp}$  of such a local minima  $\mathbf{f}_H^0$ . From Eq. (A.6),

$$\begin{aligned} \bar{J}(\mathbf{f}_H^0 + \Delta\mathbf{f}) &= \bar{J}(\mathbf{f}^0 + \Delta\mathbf{f}_H) + 2\|\Delta\mathbf{f}_{H^\perp}\|_R^4 \\ &\quad + 2\|\Delta\mathbf{f}_{H^\perp}\|_R^2(2\|\mathbf{f}_H^0 + \Delta\mathbf{f}_H\|_R^2 - r) \quad (\text{A.7}) \\ &> \bar{J}(\mathbf{f}^0) + 2\|\Delta\mathbf{f}_{H^\perp}\|_R^4 + 2\|\Delta\mathbf{f}_{H^\perp}\|_R^2(2\|\mathbf{f}_H^0\|_R^2 \\ &\quad + \|\Delta\mathbf{f}_H\|_R^2 - r). \quad (\text{A.8}) \end{aligned}$$

Because  $\|\mathbf{f}_H^0\|_R^2 > r/2$ , there exists a small neighborhood around  $\mathbf{0}$  such that the last term in the above inequality is positive. Therefore,  $\bar{J}(\mathbf{f}_H^0 + \Delta\mathbf{f}) > \bar{J}(\mathbf{f}_H^0)$ .

To prove case 3, we consider the case when the first  $M_+$  sources are super-Gaussian, i.e.,  $\pi_i > 0$  for  $i = 1, \dots, M_+$  and zero otherwise. Let  $\mathbf{f}^0 \notin \mathcal{C}_{H^\perp} \cup \mathcal{C}_{H^\varepsilon}$  be a local minimum. Because of (A1), there exists an arbitrary small  $\varepsilon > 0$  and  $\Delta\mathbf{f}_H$  such that

$$(\mathbf{f}^0 + \Delta\mathbf{f}_H)^H \mathbf{H} = (1 - \varepsilon)\mathbf{f}^0 H.$$

Since  $\sigma^2 > 0$ , choosing  $\Delta\mathbf{f}_{H^\perp}$  such that  $\|\Delta\mathbf{f}_{H^\perp}\|^2 = (1 - (1 - \varepsilon)^2)\|\mathbf{f}^0 + \Delta\mathbf{f}_H\|_R^2$ , we have

$$\|\mathbf{f}^0\|_R^2 = \|\mathbf{f}^0 + \Delta\mathbf{f}\|_R^2.$$

$$\sum_{i=1}^{M_+} \pi_i |e_i^H \mathbf{H}^H (\mathbf{f}^0 + \Delta\mathbf{f})|^4 < \sum_{i=1}^{M_+} \pi_i |e_i^H \mathbf{H}^H \mathbf{f}^0|^4.$$

This contradicts that  $\mathbf{f}^0$  is a local minimum. Now consider any local minimum  $\mathbf{f}^0$ , then

$$\begin{aligned} \bar{J}(\mathbf{f}^0) &= 2\|\mathbf{f}^0\|_R^4 - 2r\|\mathbf{f}^0\|_R^2 + \sum_{m=1}^{M_+} \pi_m (|e_m^H \mathbf{H}^H \mathbf{f}^0|^4 + r^2) \\ &= 2\|\mathbf{f}^0\|_R^4 - 2r\|\mathbf{f}^0\|_R^2. \end{aligned}$$

It is easy to see that  $\|\mathbf{f}^0\|_R^2 = r/2$ .

## Appendix B. Proof of Theorem 2

For the noiseless case,  $\Phi = \mathbf{I}_M$ . The CM cost function is given by

$$J(\mathbf{q}) = 2\|\mathbf{q}\|_2^4 - 2r\|\mathbf{q}\|_2^2 + \sum_{m=1}^M \pi_m |q_m|^4 + r^2. \quad (\text{B.1})$$

The gradient of  $J(\mathbf{q})$  is

$$\nabla_{\mathbf{q}^*} J(\mathbf{q}) = 4\|\mathbf{q}\|^2 \mathbf{q} - 2r\mathbf{q} + 2\mathbf{Q}\mathbf{q} = 2(2\|\mathbf{q}\|^2 \mathbf{I}_M - r\mathbf{I}_M + \mathbf{Q})\mathbf{q}, \quad (\text{B.2})$$

where,

$$\mathbf{Q} = \text{diag}\{\pi_1 |q_1|^2, \dots, \pi_M |q_M|^2\}.$$

The stationary points are given by

$$2\|\mathbf{q}\|^2 - r + \pi_m |q_m|^2 = 0 \text{ or } q_m = 0, \quad m = 1, 2, \dots, M. \quad (\text{B.3})$$

This implies that when there are  $K$  non-zero entries  $q_{m_j}, j = 1, 2, \dots, K$ , in a stationary point  $\mathbf{q}$ , we have three possibilities:

S0:  $\mathbf{q} = \mathbf{0}$ .

S1:  $\pi_{m_i}\pi_{m_j} > 0, \forall i, j = 1, 2, \dots, K$ ,

$$|q_{m_j}|^2 = \frac{r}{\pi_{m_j}(2\eta_K + 1)}, \eta_K \triangleq \sum_{j=1}^K \frac{1}{\pi_{m_j}}, \quad (\text{B.4})$$

$$\|\mathbf{q}\|^2 = \frac{\eta_K}{2\eta_K + 1}r. \quad (\text{B.5})$$

S2:  $\pi_{m_j} = 0, j = 1, 2, \dots, K$ ,

$$\|\mathbf{q}\|^2 = \frac{r}{2}. \quad (\text{B.6})$$

Next we will check the stationary points and obtain the local minima of  $J(\mathbf{q})$ .

The Hessian matrix of Eq. (B.1) is

$$\nabla_{\mathbf{q}}\nabla_{\mathbf{q}^*}J(\mathbf{q}) = 4\|\mathbf{q}\|^2\mathbf{I} + 4\mathbf{q}\mathbf{q}^H - 2r\mathbf{I} + 4\mathbf{Q}, \quad (\text{B.7})$$

which, according to  $K$ , has the following three forms at the stationary points defined in S1:

(i) when  $K = M$ ,

$$\nabla_{\mathbf{q}}\nabla_{\mathbf{q}^*}J(\mathbf{q}) = 4\mathbf{q}\mathbf{q}^H + \frac{2r}{2\eta_M + 1}\mathbf{I}, \quad (\text{B.8})$$

(ii) when  $1 < K < M$ ,

$$\nabla_{\mathbf{q}}\nabla_{\mathbf{q}^*}J(\mathbf{q}) = \begin{bmatrix} -2r & & & & & & & & \\ & \frac{2r}{2\eta_K + 1} & & & & & & & \\ & & \ddots & & & & & & \\ & & & \frac{2rr_{m_1}}{\pi_{m_1}(2\eta_K + 1)} & & & & & \\ & & & & \ddots & & & & \\ & & & & & 4q_{m_1}q_{m_2}^* & & * & \\ & & & & & & \ddots & & \\ & & & & & & & \frac{2rr_{m_2}}{\pi_{m_2}(2\eta_K + 1)} & \\ & & & & & * & & & \ddots & \\ & & & & & & & & & -2r \\ & & & & & & & & & \frac{2r}{2\eta_K + 1} \end{bmatrix}; \quad (\text{B.9})$$

(iii) when  $K = 1$ ,

$$\nabla_{\mathbf{q}}\nabla_{\mathbf{q}^*}J(\mathbf{q}) = \begin{bmatrix} \frac{2r(2-r_{m_1})}{r_{m_1}} & & & & & & & & \\ & \ddots & & & & & & & \\ & & \frac{2r(2-r_{m_1})}{r_{m_1}} & & & & & & \\ & & & 2r & & & & & \\ & & & & \frac{2r(2-r_{m_1})}{r_{m_1}} & & & & \\ & & & & & \ddots & & & \\ & & & & & & \frac{2r(2-r_{m_1})}{r_{m_1}} & & \\ & & & & & & & & \frac{2r(2-r_{m_1})}{r_{m_1}} \end{bmatrix}. \quad (\text{B.10})$$

Since  $\nabla_q \nabla_{q^*} J(\mathbf{q})|_{\mathbf{q}=\mathbf{0}} = -2r\mathbf{I} < 0$ ,  $\mathbf{q} = \mathbf{0}$  is a local maximum.

1. First, we show case 1 ( $M_- > 0$ ): If all sources corresponding to non-zero entries of  $\mathbf{q}$  are super-Gaussian, i.e.,  $r_{m_j} > 2$ ,

$$\frac{-2r}{2\eta_K + 1} = \frac{2r(2 - r_{m_j})}{\pi_{m_j}(2\eta_K + 1)} < 0, \quad (\text{B.11})$$

$$\left| \begin{array}{cc} \frac{2rr_{m_1}}{\pi_{m_1}(2\eta_K + 1)} & 4q_{m_1}q_{m_2}^* \\ 4q_{m_1}^*q_{m_2} & \frac{2rr_{m_2}}{\pi_{m_2}(2\eta_K + 1)} \end{array} \right| = \frac{4r^2(r_{m_1}r_{m_2} - 4)}{\pi_{m_1}\pi_{m_2}(2\eta_K + 1)^2} > 0. \quad (\text{B.12})$$

Hence  $\nabla_q \nabla_{q^*} J(\mathbf{q})$  is indefinite and the stationary point is not stable. If all sources corresponding to non-zero entries of  $\mathbf{q}$  are sub-Gaussian, i.e.,  $r_{m_i} < 2$ , then

$$\frac{-2r}{2\eta_K + 1} > 0 \quad (\text{B.13})$$

and

$$\left| \begin{array}{cc} \frac{2rr_{m_1}}{\pi_{m_1}(2\eta_K + 1)} & 4q_{m_1}q_{m_2}^* \\ 4q_{m_1}^*q_{m_2} & \frac{2rr_{m_2}}{\pi_{m_2}(2\eta_K + 1)} \end{array} \right| < 0. \quad (\text{B.14})$$

Eqs. (B.13) and (B.14) again imply that  $\nabla_q \nabla_{q^*} J(\mathbf{q})$  of Eq. (B.9) is indefinite; therefore, there is no local minimum when  $2 \leq K \leq M$  and  $\pi_{m_j} \neq 0$ .

When  $r_{m_j} = 2$ ,  $j = 1, 2, \dots, K$ , consider a stationary point  $\mathbf{q}$  in S2,

$$\mathbf{q} = \begin{bmatrix} \mathbf{q}_- \\ \mathbf{q}_0 \\ \mathbf{q}_+ \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{q}_0 \\ \mathbf{0} \end{bmatrix}, \quad \|\mathbf{q}\|^2 = \|\mathbf{q}_0\|^2 = \frac{r}{2}. \quad (\text{B.15})$$

If  $M_- > 0$ , there exists a small enough

$$\Delta \mathbf{q} = \begin{bmatrix} \Delta \mathbf{q}_- \\ \Delta \mathbf{q}_0 \\ \mathbf{0} \end{bmatrix}$$

such that

$$\|\mathbf{q} + \Delta \mathbf{q}\|^2 = \frac{r}{2}.$$

Thus  $J(\mathbf{q} + \Delta \mathbf{q}) - J(\mathbf{q}) = \sum_{m=1}^{M_-} \pi_m |\Delta q_m|^4 < 0$ , which means  $\mathbf{q}$  given in S2 are not local minima. When  $K = 1$ , we have

$$\frac{2r(2 - r_{m_1})}{r_{m_1}} > 0 \quad \text{iff} \quad r_{m_1} < 2.$$

Therefore,  $\mathbf{q} = q_m \mathbf{e}_m$  is a local minimum iff  $r_m < 2$ . This proves

$$A = \left\{ q_m \mathbf{e}_m \mid |q_m|^2 = \frac{r}{r_m}, 1 \leq m \leq M_- \right\}. \quad (\text{B.16})$$



2. Now let us consider case 2 ( $M_+ = M$ ): From  $r_m > 2$ ,  $m = 1, 2, \dots, M$ , it is easy to see that Eqs. (B.10) and (B.9) are indefinite and Eq. (B.8) is positive definite, i.e.,

$$A = \left\{ \mathbf{q} = (q_1, q_2, \dots, q_M)^T, |q_m|^2 = \frac{r}{\pi_m(2\eta_M + 1)} \right\}.$$

3. Finally, assume  $M_- = 0, M_0 > 0$ : Considering

$$\mathbf{q} = \begin{bmatrix} \mathbf{q}_- \\ \mathbf{q}_0 \\ \mathbf{q}_+ \end{bmatrix}$$

in Eq. (B.15), if  $M_- = 0$ , for any  $\Delta \mathbf{q} \neq 0$ ,  $J(\mathbf{q} + \Delta \mathbf{q}) - J(\mathbf{q}) > 0$ , which implies that  $\mathbf{q}$  is a local minimum.

### Appendix C. Proof of Theorem 3

The proof has four steps: (i) express the CM cost function using the extra MSE  $\delta$  defined in Eq. (21) and the receiver gain  $\theta$  and  $\theta_w$ ; (ii) give the reference  $\mathbf{q}_r$ ; (iii) prove there is a local minimum in  $\mathcal{B}(\mathbf{q}_w, \delta_U, \theta_L, \theta_U)$ ; and (iv) verify that when noise variance goes to 0, CM receiver goes to the Wiener receiver.

1. Suppose that  $\mathbf{q}_w = \theta_w \begin{pmatrix} 1 \\ \mathbf{q}_{w1} \end{pmatrix}$  is the Wiener receiver for  $s_1$ . For any receiver

$$\mathbf{q} = \theta \begin{pmatrix} 1 \\ \mathbf{q}_1 \end{pmatrix},$$

it can be shown [13,14] that

$$E\{y^2\} = \|\mathbf{q}\|_{\Phi}^2 = |\theta|^2 \left( \frac{1}{\theta_w} + \delta^2 \right). \quad (\text{C.1})$$

According to Eq. (C.1), the CM cost function  $J(\mathbf{q})$  is then given by

$$J(\mathbf{q}) = 2 \left( \delta^2 + \frac{1}{\theta_w} \right)^2 |\theta|^4 - 2r \left( \delta^2 + \frac{1}{\theta_w} \right) |\theta|^2 + (\pi_1 + \pi_{\min} \|\mathbf{I}_1 \mathbf{q}_1\|_4^4) |\theta|^4 + r^2, \quad (\text{C.2})$$

where  $\mathbf{I}_1$  and  $\pi_{\min}$  are defined as in Eqs. (26) and (25), respectively.

In the following, Eq. (C.2) will be used to locate the local minima of  $J(\mathbf{q})$ .

2. Choose reference  $\mathbf{q}_r \triangleq \theta_r \begin{pmatrix} 1 \\ \mathbf{q}_{r1} \end{pmatrix}$  with the minimum CM cost:

$$J \left( \theta \begin{pmatrix} 1 \\ \mathbf{q}_{w1} \end{pmatrix} \right) = \frac{2}{\theta_w^2} |\theta|^4 - \frac{2r}{\theta_w} |\theta|^2 + (\pi_1 + \pi_{\min} \|\mathbf{I}_1 \mathbf{q}_{w1}\|_4^4) |\theta|^4 + r^2. \quad (\text{C.3})$$

From

$$\frac{\partial J}{\partial |\theta|^2} = 2 \left[ \frac{2}{\theta_w^2} + (\pi_1 + \pi_{\min} \|\mathbf{I}_1 \mathbf{q}_{w1}\|_4^4) \right] |\theta|^2 - \frac{2r}{\theta_w} = 0, \quad (\text{C.4})$$

we get

$$|\theta_r|^2 = \frac{r\theta_w}{2 + \theta_w^2(\pi_1 + \pi_{\min}\|\mathbf{P}_1\mathbf{q}_{w1}\|_4^4)}. \quad (\text{C.5})$$

Thus,

$$\begin{aligned} J(\mathbf{q}_r) &= \frac{2}{\theta_w^2} \frac{r^2\theta_w^2}{[2 + \theta_w^2(\pi_1 + \pi_{\min}\|\mathbf{P}_1\mathbf{q}_{w1}\|_4^4)]^2} \\ &\quad - \frac{2r}{\theta_w} \frac{r\theta_w}{2 + \theta_w^2(\pi_1 + \pi_{\min}\|\mathbf{P}_1\mathbf{q}_{w1}\|_4^4)} + (\pi_1 + \pi_{\min}\|\mathbf{P}_1\mathbf{q}_{w1}\|_4^4) \frac{r^2\theta_w^2}{[2 + \theta_w^2(\pi_1 + \pi_{\min}\|\mathbf{P}_1\mathbf{q}_{w1}\|_4^4)]^2} + r^2 \\ &= \frac{r^2[2 + \theta_w^2(\pi_1 + \pi_{\min}\|\mathbf{P}_1\mathbf{q}_{w1}\|_4^4)]}{[2 + \theta_w^2(\pi_1 + \pi_{\min}\|\mathbf{P}_1\mathbf{q}_{w1}\|_4^4)]^2} - \frac{2r^2}{2 + \theta_w^2(\pi_1 + \pi_{\min}\|\mathbf{P}_1\mathbf{q}_{w1}\|_4^4)} + r^2 \\ &= r^2 - \frac{r^2}{2 + \theta_w^2(\pi_1 + \pi_{\min}\|\mathbf{P}_1\mathbf{q}_{w1}\|_4^4)} \end{aligned} \quad (\text{C.6})$$

$$= r^2 - r \frac{|\theta_r|^2}{\theta_w}. \quad (\text{C.7})$$

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

$$\begin{aligned} \|\mathbf{P}_1\mathbf{q}_1\|_4 &= \|\mathbf{P}_1\mathbf{q}_1 - \mathbf{P}_1\mathbf{q}_{w1} + \mathbf{P}_1\mathbf{q}_{w1}\|_4 \\ &\leq \|\mathbf{P}_1(\mathbf{q}_1 - \mathbf{q}_{w1})\|_4 + \|\mathbf{P}_1\mathbf{q}_{w1}\|_4 \\ &< \|\mathbf{q}_1 - \mathbf{q}_{w1}\|_C + \|\mathbf{P}_1\mathbf{q}_{w1}\|_4 \\ &= \delta + \|\mathbf{P}_1\mathbf{q}_{w1}\|_4. \end{aligned}$$

Therefore,

$$\begin{aligned} J(\mathbf{q}) - J(\mathbf{q}_r) &= \left[ 2\left(\delta^2 + \frac{1}{\theta_w}\right)^2 + \pi_1 + \pi_{\min}\|\mathbf{P}_1\mathbf{q}_1\|_4^4 \right] |\theta|^4 \\ &\quad - 2r\left(\delta^2 + \frac{1}{\theta_w}\right) |\theta|^2 + r^2 + \frac{r^2}{2 + \theta_w^2(\pi_1 + \pi_{\min}\|\mathbf{P}_1\mathbf{q}_{w1}\|_4^4)} \\ &\quad - r^2 > \left[ 2\left(\delta^2 + \frac{1}{\theta_w}\right)^2 + \pi_1 + \pi_{\min}(\delta + \|\mathbf{P}_1\mathbf{q}_{w1}\|_4)^4 \right] |\theta|^4 \\ &\quad - 2r\left(\delta^2 + \frac{1}{\theta_w}\right) |\theta|^2 + \frac{r^2}{2 + \theta_w^2(\pi_1 + \pi_{\min}\|\mathbf{P}_1\mathbf{q}_{w1}\|_4^4)} = c_2(\delta)|\theta|^4 + c_1(\delta)|\theta|^2 + c_0. \end{aligned} \quad (\text{C.8})$$

3. To show that there is CM receiver in  $\mathcal{B}(\mathbf{q}_w, \delta_U, \theta_L, \theta_U)$ , we need to show the CM cost of every point on the boundary of  $\mathcal{B}(\mathbf{q}_w, \delta_U, \theta_L, \theta_U)$  is greater than that of reference  $\mathbf{q}_r$  and the reference is inside  $\mathcal{B}(\mathbf{q}_w, \delta_U, \theta_L, \theta_U)$ .  
 (a) Consider all points on the peripheral surface  $\mathcal{S}_p = \{\mathbf{q}: \|\mathbf{q}_1 - \mathbf{q}_{w1}\|_C = \delta_U, \theta_L \leq \theta \leq \theta_U\}$  shown in Fig. 10. First we need to show  $c_2(\delta) > 0, \forall \delta \in (0, \|\mathbf{q}_{w1}\|_2)$ .

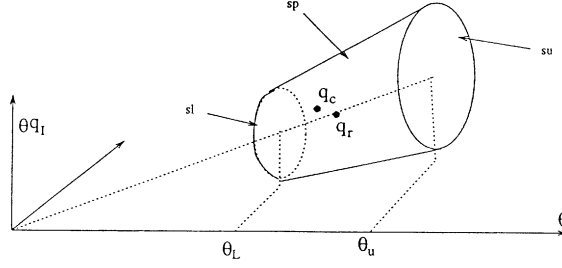


Fig. 10. Three surfaces of the sliced cone.

According to Eq. (31),

$$c_2(\delta) = 2\left(\delta^2 + \frac{1}{\theta_w}\right)^2 + \pi_1 + \pi_{\min}(\delta + \|\mathbf{P}_1 \mathbf{q}_{w1}\|_4)^4 = (2 + \pi_{\min})\delta^4 + 4\pi_{\min}\|\mathbf{P}_1 \mathbf{q}_{w1}\|_4 \delta^3 + \left(\frac{4}{\theta_w} + 6\pi_{\min}\|\mathbf{P}_1 \mathbf{q}_{w1}\|_4^2\right)\delta^2 + 4\pi_{\min}\|\mathbf{P}_1 \mathbf{q}_{w1}\|_4^3 \delta + \frac{2}{\theta_w^2} + \pi_1 + \pi_{\min}\|\mathbf{P}_1 \mathbf{q}_{w1}\|_4^4. \quad (\text{C.9})$$

From  $1 - \theta_w \geq (1 - \theta_w)^2 + \theta_w^2 \|\mathbf{q}_{w1}\|_2^2$  [14], we have  $\|\mathbf{q}_{w1}\|_2^2 \leq (1 - \theta_w)/\theta_w$ . Therefore,

$$\|\mathbf{P}_1 \mathbf{q}_{w1}\|_4 < \|\mathbf{q}_{w1}\|_2 < \sqrt{\frac{1 - \theta_w}{\theta_w}}. \quad (\text{C.10})$$

Substituting Eq. (C.10) into Eq. (C.9), it can be obtained

$$\begin{aligned} c_2(\delta) &> (2 + \pi_{\min})\delta^4 + 4\pi_{\min}\|\mathbf{q}_{w1}\|_2^4 + \left(\frac{4}{\theta_w} + 6\pi_{\min}\|\mathbf{q}_{w1}\|_2^2\right)\delta^2 + 4\pi_{\min}\|\mathbf{q}_{w1}\|_2^4 + \frac{2}{\theta_w^2} + \pi_1 \\ &+ \pi_{\min}\|\mathbf{q}_{w1}\|_2^4 > (2 + \pi_{\min})\delta^4 + \left(\frac{4}{\theta_w} + 6\pi_{\min}\frac{1 - \theta_w}{\theta_w}\right)\delta^2 + \frac{2}{\theta_w^2} + \pi_1 + 9\pi_{\min}\left(\frac{1 - \theta_w}{\theta_w}\right)^2 \\ &= (2 + \pi_{\min})\delta^4 + \frac{4 + 6\pi_{\min} - 6\pi_{\min}\theta_w}{\theta_w}\delta^2 + \frac{1}{\theta_w^2}[\pi_1 + 9\pi_{\min}\theta_w^2 - 18\pi_{\min}\theta_w + (2 + 9\pi_{\min})]. \end{aligned} \quad (\text{C.11})$$

Now we want to find the condition for  $\theta_w$  in which  $c_2(\delta) > 0$  for any  $\delta \in (0, \|\mathbf{q}_{w1}\|_2)$ . There are three terms in Eq. (C.11). If

$$4 + 6\pi_{\min} - 6\pi_{\min}\theta_w > 0 \quad (\text{C.12})$$

and

$$(\pi_1 + 9\pi_{\min})\theta_w^2 - 18\pi_{\min}\theta_w + (2 + 9\pi_{\min}) > 0, \quad (\text{C.13})$$

then  $c_2(\delta) > 0$ .

Eq. (C.12) is equivalent to

$$\theta_w > 1 + \frac{2}{3\pi_{\min}}. \quad (\text{C.14})$$

The left-hand side of Eq. (C.13) is a polynomial of  $\theta_w$ . Define

$$u(\pi_1, \pi_{\min}) \triangleq (18\pi_{\min})^2 - 4(\pi_1 + 9\pi_{\min})(2 + 9\pi_{\min}), \quad (\text{C.15})$$

then Eq. (C.13) is equivalent to

$$\frac{18\pi + \sqrt{u(\pi_1, \pi_{\min})}}{2(\pi_1 + 9\pi_{\min})} < \theta_w < \frac{18\pi_{\min} - \sqrt{u(\pi_1, \pi_{\min})}}{2(\pi_1 + 9\pi_{\min})}. \quad (\text{C.16})$$

It is easy to see

$$u(\pi_1, \pi_{\min}) = 4[-2\pi_1 - 9\pi_{\min}(2 + \pi_1)] > 4[-2\pi_1] > 4\pi_1^2,$$

hence,  $(18\pi_{\min} - \sqrt{u(\pi_1, \pi_{\min})})/2(\pi_1 + 9\pi_{\min}) > 1$ , which means  $\theta_w < (18\pi_{\min} - \sqrt{u(\pi_1, \pi_{\min})})/2(\pi_1 + 9\pi_{\min})$  is always true.

Now we have shown that  $c_2(\delta) > 0$ ,  $\forall \delta \in (0, \|\mathbf{q}_{wl}\|_2)$ , when Eq. (32) holds.

Also we know  $0 < \delta_U < \|\mathbf{q}_{wl}\|_2$ ; therefore,

$$J(\mathbf{q}) - J(\mathbf{q}_r) > c_2(\delta_U)|\theta|^4 + c_1(\delta_U)|\theta|^2 + c_0 \geq 0, \quad \forall \mathbf{q} \in \mathcal{S}_p. \quad (\text{C.17})$$

(b) Now we check the points on the upper surface  $\mathcal{S}_U$  defined by  $\theta = \theta_U, 0 \leq \delta \leq \delta_U$ . Since

$$|\theta|^2 = (\theta_U)^2 \geq \frac{-c_1(\delta) + \sqrt{c_1(\delta)^2 - 4c_2(\delta)c_0}}{2c_2(\delta)}, \quad (\text{C.18})$$

and  $c_2(\delta) > 0$ , then the polynomial  $c_2(\delta)|\theta|^4 + c_1(\delta)|\theta|^2 + c_0 \geq 0$ . Thus,  $J(\mathbf{q}) - J(\mathbf{q}_r) > 0$ . Similarly,  $J(\mathbf{q}) - J(\mathbf{q}_r) > 0 \forall \mathbf{q} \in \mathcal{S}_L$ .

(c) Finally, we verify that  $(\theta_L)^2 < |\theta_r|^2 < (\theta_U)^2$ . From  $|\theta_r|^2 = -c_1(0)/2c_2(0)$  and Eqs. (27) and (28), it can be obtained easily.

4. In the end, we examine the case as the noise variance  $\sigma \rightarrow 0$ . From Eqs. (29)–(31), we have

$$\begin{aligned} \lim_{\sigma \rightarrow 0} c_0 &= \lim_{\sigma \rightarrow 0} \frac{r^2}{2 + \pi_1 \theta_w^2 + \pi_{\min} \theta_w^2 \|\mathbf{P}\mathbf{P}\mathbf{q}_{wl}\|_4^4} = r, \\ \lim_{\sigma \rightarrow 0} c_1(\delta) &= -2r(\delta^2 + 1), \\ \lim_{\sigma \rightarrow 0} c_2(\delta) &= 2(\delta^2 + 1)^2 + \pi_1 + \pi_{\min} \delta^4, \\ \lim_{\sigma \rightarrow 0} D(\delta) &= 4r^2(\delta^2 + 1)^2 - 4r[(2(\delta^2 + 1)^2 + \pi_1 + \pi_{\min} \delta^4)], = 4r(\pi_1 - \pi_{\min})\delta^4 + 8r\pi_1 \delta^2. \end{aligned}$$

So

$$\lim_{\sigma^2 \rightarrow 0} \delta_U = 0$$

and

$$\lim_{\sigma^2 \rightarrow 0} \theta_L = \lim_{\sigma \rightarrow 0} \theta_U = 1 = \theta_w.$$

#### Appendix D. Proof of Theorem 4

1. It has already been shown that  $\|\mathbf{f}_c\|_R^2 > r/2$  in Section 3. Define a function for any  $\mathbf{f}$  such that  $\|\mathbf{f}\|_R = 1$ ,

$$\phi(k, \mathbf{f}) \triangleq \bar{J}(\sqrt{k}\mathbf{f}) = (2 + \pi_{\min} \|\mathbf{P}\mathbf{H}^H \mathbf{f}\|_4^4)k^2 - 2rk + r^2, \quad (\text{D.1})$$

where

$$\mathbf{\Pi} \triangleq \text{diag} \left\{ \left( \frac{\pi_1}{\pi_{\min}} \right)^{1/4}, \dots, \left( \frac{\pi_M}{\pi_{\min}} \right)^{1/4} \right\},$$

then

$$\frac{\partial \phi(k, \mathbf{f})}{\partial k} = 2(2 + \pi_{\min} \|\mathbf{\Pi} \mathbf{H}^H \mathbf{f}\|_4^4)k - 2r. \quad (\text{D.2})$$

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

$$k_0 = \frac{r}{2 + \pi_{\min} \|\mathbf{\Pi} \mathbf{H}^H \mathbf{f}\|_4^4}. \quad (\text{D.3})$$

Since  $\pi_{\min} < 0$ , and  $\|\mathbf{\Pi} \mathbf{H}^H \mathbf{f}\|_4^4 \leq \|\mathbf{H}^H \mathbf{f}\|_2^4 < \|\mathbf{f}\|_R^4 = 1$ ,

$$\|\mathbf{f}_c\|_R^2 = k_0 < \frac{r}{\pi_{\min} + 2} = \frac{r}{r_{\min}}.$$

2. Now consider a local minimum of  $J(\mathbf{q})$  in  $\mathcal{B}(\mathbf{q}_w, \delta_U, \theta_L, \theta_U)$ .

(a) Any point in  $\mathcal{B}(\mathbf{q}_w, \delta_U, \theta_L, \theta_U)$  can be represented by  $\sqrt{k}\mathbf{q}$ , where  $\|\mathbf{q}\|_{\Phi}^2 = \|\mathbf{q}_w\|_{\Phi}^2 = \theta_w$  and  $0 \leq k < \infty$ . Define

$$\phi(k, \mathbf{q}) \triangleq J(\sqrt{k}\mathbf{q}) = 2\theta_w^2 k^2 - 2r\theta_w k + \pi_{\min} \|\mathbf{\Pi} \mathbf{q}\|_4^4 k^2 + r^2. \quad (\text{D.4})$$

From  $(d/dk)\phi(k, \mathbf{q}) = 0$ , the minimum  $\sqrt{k}\mathbf{q}$  must have

$$k = \frac{r\theta_w}{2\theta_w^2 + \pi_{\min} \|\mathbf{\Pi} \mathbf{q}\|_4^4}. \quad (\text{D.5})$$

If  $k < 1$ , then  $\|\mathbf{q}_c\|_{\Phi}^2 < \|\mathbf{q}_w\|_{\Phi}^2$ . In order to get the condition for  $k < 1$ , we adopt the following two steps:

(i) Bound  $\|\mathbf{\Pi} \mathbf{q}\|_4^4$  by  $|\theta|$  and  $\theta_w$ : Since

$$\|\mathbf{\Pi} \mathbf{q}\|_4^4 < \|\mathbf{q}\|_4^4 = |\theta|^4 + \|\theta_{\mathbf{q}_1}\|_4^4 \leq |\theta|^4 + \|\theta_{\mathbf{q}_1}\|_2^4 \leq |\theta|^4 + (\mathbf{q}^T \Phi \mathbf{q} - |\theta|^2)^2 = |\theta|^4 + (\theta_w - |\theta|^2)^2,$$

then

$$k < \frac{r\theta_w}{2\theta_w^2 + \pi_{\min}(|\theta|^4 + (\theta_w - |\theta|^2)^2)} = \frac{r\theta_w}{2\pi_{\min}|\theta|^4 - 2\pi_{\min}\theta_w|\theta|^2 + r_{\min}\theta_w^2}. \quad (\text{D.6})$$

Define a polynomial of  $|\theta|$  as

$$p(|\theta|) \triangleq 2\pi_{\min}|\theta|^4 - 2\pi_{\min}\theta_w|\theta|^2 + r_{\min}\theta_w^2 - r\theta_w, \quad (\text{D.7})$$

then  $p(|\theta|) > 0$  results in  $k > 1$ . It is easy to see that if

$$\frac{2\pi_{\min}\theta_w + \sqrt{\Delta}}{4\pi_{\min}} < |\theta|^2 < \frac{2\pi_{\min}\theta_w - \sqrt{\Delta}}{4\pi_{\min}}, \quad (\text{D.8})$$

then  $p(|\theta|) > 0$ , provided that

$$\theta_w > \frac{2r}{2 + r_{\min}}, \quad (\text{D.9})$$

which means  $\Delta = 4\pi_{\min}\theta_w(\pi_{\min}\theta_w - 2r_{\min}\theta_w + 2r) > 0$ .

(ii) Bound  $|\theta|$  by  $\theta_w$ : According to  $\mathbf{q}^T \Phi \mathbf{q} = |\theta|^2(\delta^2 + 1/\theta_w) = \theta_w$ , and  $\delta^2 < \|\mathbf{q}_{w1}\|_2^2 < (1 - \theta_w)/\theta_w$ , we know

$$\frac{\theta_w^2}{2 - \theta_w} \leq |\theta|^2 \leq \theta_w^2.$$

It can be shown that

$$2\pi_{\min}\theta_w^3 - 2\pi_{\min}\theta_w^2 + r_{\min}\theta_w^2 - r\theta_w \geq 0 \Rightarrow \theta_w^2 < \frac{2\pi_{\min}\theta_w - \sqrt{\Delta}}{4\pi_{\min}}$$

and

$$\theta_w > \frac{2}{3} \Rightarrow \frac{\theta_w^2}{2 - \theta_w} > \frac{2\pi_{\min}\theta_w + \sqrt{\Delta}}{4\pi_{\min}}.$$

Therefore,  $\|\mathbf{q}_c\|_{\Phi}^2 < \|\mathbf{q}_w\|_{\Phi}^2 < 1$  and  $|\theta_c| \leq \theta_w$ , when Eqs. (36) and (37) hold.

(b) Prove by contradiction. Suppose that  $\|\mathbf{q}_{c1}\|_4 < \|\mathbf{q}_{w1}\|_4$ . Let

$$\hat{\mathbf{q}}_c \triangleq \theta_c \begin{pmatrix} 1 \\ \mathbf{q}_{w1} \end{pmatrix}. \quad (\text{D.10})$$

Obviously,  $\hat{\mathbf{q}}_c \in \mathcal{B}_p(\mathbf{q}_w, \delta_U, \theta_L, \theta_U)$ , and  $\|\hat{\mathbf{q}}_c\|_4 > \|\mathbf{q}_c\|_4$ . According to Eq. (C.1),

$$\|\hat{\mathbf{q}}_c\|_{\Phi}^2 = \frac{\theta_c^2}{\theta_w} \leq \left( \delta_c^2 + \frac{1}{\theta_w} \right) \theta_c^2 \leq \|\mathbf{q}_c\|_{\Phi}^2. \quad (\text{D.11})$$

Also we have known  $\|\hat{\mathbf{q}}_c\|_{\Phi}^2 > r/2$ , therefore

$$J(\hat{\mathbf{q}}_c) < J(\mathbf{q}_c) \quad (\text{D.12})$$

which contradicts that  $\mathbf{q}_c$  is the CMA minimum in  $\mathcal{B}(\mathbf{q}_w, \delta_U, \theta_L, \theta_U)$ .

## Appendix E. Proof of Theorem 5

$$\begin{aligned} \Delta^{\mathcal{E}} &= J_w \left( \theta_c \begin{pmatrix} 1 \\ \mathbf{q}_{c1} \end{pmatrix} \right) - J_w(\mathbf{q}_w) = 1 - 2\text{Re}\{\theta_c\} + |\theta_c|^2 \left( \delta_c^2 + \frac{1}{\theta_w} \right) - (1 - \theta_w) \\ &= \frac{|\theta_c|^2 - 2\text{Re}\{\theta_c\}\theta_w + \theta_w^2}{\theta_w} + |\theta_c|^2 \delta_c^2. \end{aligned} \quad (\text{E.1})$$

If  $\theta_c$  is real,

$$\Delta^{\mathcal{E}} = \frac{(\theta_c - \theta_w)^2}{\theta_w} + \theta_c^2 \delta_c^2. \quad (\text{E.2})$$

Thus, we obtain the bounds in Eq. (41):

$$\widehat{\Delta}^{\mathcal{E}} = \frac{(\theta_r - \theta_w)^2}{\theta_w} = \left( \frac{\theta_r}{\sqrt{\theta_w}} - \sqrt{\theta_w} \right)^2. \quad (\text{E.3})$$

According to Eq. (C.5),

$$\frac{\theta_r}{\sqrt{\theta_w}} = \frac{(r\theta_w)^{1/2}}{\theta_w^{1/2}(2 + \pi_1\theta_w^2 + \pi_{\min}\theta_w^2\|\mathbf{P}_1\mathbf{q}_{w1}\|_4^4)^{1/2}}, \quad (\text{E.4})$$

hence,

$$\widehat{\Delta\mathcal{E}} = (\sqrt{r/(2 + \pi_1\theta_w^2 + \pi_{\min}\theta_w^2\|\mathbf{P}_1\mathbf{q}_{w1}\|_4^4)} - \sqrt{\theta_w})^2.$$

From Eq. 100, we have

$$\begin{aligned} \frac{\theta_r}{\sqrt{\theta_w}} &= \left(1 - \frac{\pi_1}{r} \left(1 - \theta_w^2 - \frac{\pi_{\min}\theta_w^2\|\mathbf{P}_1\mathbf{q}_{w1}\|_4^4}{\pi_1}\right)\right)^{-1/2} \\ &= 1 + \frac{\pi_1}{2r} \left(1 - \theta_w^2 - \theta_w^2 \frac{\pi_{\min}\|\mathbf{P}_1\mathbf{q}_{w1}\|_4^4}{\pi_1}\right) + \mathcal{O}\left(1 - \theta_w^2 - \theta_w^2 \frac{\pi_{\min}\|\mathbf{P}_1\mathbf{q}_{w1}\|_4^4}{\pi_1}\right). \end{aligned}$$

Since

$$1 - \theta_w^2 - \theta_w^2 \frac{\pi_{\min}\|\mathbf{P}_1\mathbf{q}_{w1}\|_4^4}{\pi_1} = (1 + \theta_w)(1 - \theta_w) - \theta_w^2 \frac{\pi_{\min}\|\mathbf{P}_1\mathbf{q}_{w1}\|_4^4}{\pi_1} = 2J_w(\mathbf{q}_w) + \mathcal{O}(J_w(\mathbf{q}_w))$$

and

$$\sqrt{\theta_w} = (1 - J_w(\mathbf{q}_w))^{1/2} = 1 - \frac{1}{2}J_w(\mathbf{q}_w) + \mathcal{O}(J_w(\mathbf{q}_w)),$$

then

$$\begin{aligned} \frac{\theta_r}{\sqrt{\theta_w}} - \sqrt{\theta_w} &= \left[1 + \frac{\pi_1}{2r} 2J_w(\mathbf{q}_w)\right] - \left[1 - \frac{1}{2}J_w(\mathbf{q}_w)\right] + \mathcal{O}(J_w(\mathbf{q}_w)) \\ &= \frac{3r - 4}{2r} J_w(\mathbf{q}_w) + \mathcal{O}(J_w(\mathbf{q}_w)), \end{aligned}$$

i.e.

$$\widehat{\Delta\mathcal{E}} = \left(\frac{3r - 4}{2r}\right)^2 J_w^2(\mathbf{q}_w) + \mathcal{O}(J_w^2(\mathbf{q}_w)). \quad (\text{E.5})$$

## References

- [1] J.M. Cioffi, G.P. Dudevoir, M.V. Eyuboglu, G.D. Forney, Jr. MMSE decision feedback equalization and coding-Part I and II, *IEEE Trans. Commun.* COM-43 (10) (October 1995) 2582–2604.
- [2] G.J. Foschini, Equalizing without altering or detecting data, *Bell Syst. Tech. J.* 64 (October 1985) 1885–1911.
- [3] D.N. Godard, Self-recovering equalization and carrier tracking in two-dimensional data communication systems, *IEEE Trans. Commun.* COM-28 (11) (November 1980) 1867–1875.
- [4] R.P. Gooch, J.D. Lundell, The CM array: An adaptive beamformer for constant modulus signals, in: *Proc. ICASSP 86 Conf.*, Tokyo, April 1986, pp. 2523–2526.
- [5] A.V. Keerthi, A. Mathur, J.J. Shynk, Cochannel signal recovery using the MUSIC algorithm and the constant modulus array, *IEEE Trans. Signal Process Lett.* 2 (October 1995) 191–194.
- [6] A.V. Keerthi, J.J. Shynk, A. Mathur, Steady-state analysis of the multistage constant modulus array, *IEEE Trans. Signal Process.* SP-44 (4) (April 1996) 948–962.
- [7] P. Lancaster, M. Tismenetsky, *The Theory of Matrices*, Academic Press, New York, NY, 1984.
- [8] J.J. Shynk, R.P. Gooch, The constant modulus array for cochannel signal copy and direction finding, *IEEE Trans. Signal Process.* SP-44 (3) (March 1996) 652–660.
- [9] J.J. Shynk, A. Mathur, A.V. Keerthi, R.P. Gooch, Convergence properties of the multistage constant modulus array for correlated sources, *IEEE Trans. Signal Process.* 45 (January 1997) 280–286.
- [10] L. Tong, H. Zeng, Channel-surfing re-initialization for the constant modulus algorithm, *IEEE Signal Process Lett.* 4 (March 1997) 85–87.
- [11] J.R. Treichler, B.G. Agee, A new approach to multipath correction of constant modulus signals, *IEEE Trans. Acoust. Speech Signal Process.* ASSP-31 (2) (April 1983) 459–472.

- [12] A. van der Veen, A. Paulraj, An analytical constant modulus algorithm, *IEEE Trans. Signal Process.* SP-44 (5) (May 1996) 1136–1155.
- [13] H. Zeng, 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, March 1996, pp. 890–894.
- [14] H. Zeng, L. Tong, C.R. Johnson, Relationships between CMA and Wiener receivers, *IEEE Trans. Inform. Theory* 44 (4) (July 1998).
- [15] H. Zeng, L. Tong, C.R. Johnson, An analysis of constant modulus receivers, *IEEE Trans. Signal Process.*, accepted.