Bounds for the MSE Performance of Constant Modulus Estimators

Philip Schniter and C. Richard Johnson, Jr.
Cornell University, Ithaca, NY 14853.

Abstract— **The constant modulus (CM) criterion has become popular in the design of blind linear estimators of sub-Gaussian i.i.d. processes transmitted through unknown linear channels in the presence of unknown additive interference. In this paper, we present an upper bound for the conditionally unbiased mean-squared error (UMSE) of CM-minimizing estimates that depends only on the source kurtoses and the UMSE of Wiener estimates. Further analysis reveals that the extra UMSE of CM estimates can be upper bounded by approximately the square of the Wiener (i.e., minimum) UMSE. Since our results hold for arbitrary linear channels and additive multi-source interference, they confirm the longstanding conjecture regarding the general MSE-robustness of CM estimates.**

I. Introduction

Consider the linear estimation problem of Fig. 1, where a desired source sequence $\{s_n^{(0)}\}$ combines linearly with K interfering sources $\{s_n^{(k)}\}$ through vector channels $\{\mathbf{h}^{(0)}(z), \dots, \mathbf{h}^{(K)}(z)\}$. Our goal is to estimate the desired source using the vector linear estimator $\mathbf{f}(z)$. The linear estimates $\{y_n\}$ which minimize the mean-squared error (MSE)

$$J_{\mathrm{m},\nu}(y_n) \;:=\; \mathrm{E}\{|y_n - s_{n-\nu}^{(0)}|^2\} \tag{1}$$

are generated by the minimum MSE (MMSE) estimator, or Wiener estimator, $\mathbf{f}_{\mathrm{m},\nu}(z)$. Specification of $\mathbf{f}_{\mathrm{m},\nu}(z)$, however, requires knowledge of the joint statistics of the observed sequence $\{\mathbf{r}_n\}$ and the desired source sequence $\{s_n^{(0)}\}$, which are typically unavailable when the channel is unknown.

When only the statistics of the observed sequence $\{\mathbf{r}_n\}$ are known, it may still be possible to estimate $\{s_n^{(0)}\}$ up to unknown magnitude and delay, i.e., $y_n = \sum_i \mathbf{f}_i^H \mathbf{r}_{n-i} \approx \alpha s_{n-\nu}^{(0)}$ for some $\alpha \in \mathbb{C}$, some $\nu \in \mathbb{Z}$, and all n. The literature refers to this problem as *blind* estimation (or blind deconvolution).

Minimization of the constant modulus (CM) criterion [1], [2] has become perhaps the most studied and implemented means of blind equalization for data communication over dispersive channels (see, e.g., [3] and the references within) and has also been used successfully as a means of blind beamforming (see, e.g., [4]). The CM criterion is defined below in terms of the estimates $\{y_n\}$ and a design parameter γ.

$$J_{\mathrm{c}}(y_n) \;:=\; \mathrm{E}\big\{\big(|y_n|^2 - \gamma\big)^2\big\}. \tag{2}$$

Supported in part by NSF grants MIP-9528363 and ECS-9811297 and Applied Signal Technology, Inc. Direct all correspondence to *schniter@ee.cornell.edu*.

The popularity of the CM criterion is usually attributed to

1. the existence of a simple adaptive algorithm ("CMA" [1], [2]) for estimation and tracking of the CM-minimizing estimator $\mathbf{f}_{\mathrm{c}}(z)$, and
2. the excellent MSE performance of CM-minimizing estimates.

The second of these two points was first conjectured in the original works [1], [2] and provides the theme for the recently-published comprehensive survey [3]. In this paper, we attempt to precisely quantify the general MSE performance of CM-minimizing estimates.

The last decade has seen a plethora of papers giving evidence for the "robustness" of CM performance in situations where the CM-minimizing (and MMSE) estimates are not perfect. Most of these studies, however, focus on *particular* features of the system model that prevent perfect estimation, such as

1. the presence of additive white Gaussian noise (AWGN) corrupting the observation (e.g., [5], [6], [7]),
2. channels that do not provide adequate diversity, (e.g., [5], [8]), or
3. estimators with an insufficient number of adjustable parameters (e.g., [9], [10]).

A notable exception is the work of Zeng et al. [11], in which an algorithm is given to bound the MSE of CM-minimizing estimates for the case of a single source transmitted through a finite-duration impulse response (FIR) linear channel in the presence of AWGN. The channel model assumed by [11] is general enough to incorporate most combinations of the three conditions above, though not as general as the multi-source model of Fig. 1. The bounding algorithm in [11] is rather involved, however, preventing a direct link between the MSE performance of CM and Wiener receivers.

The main contribution of this paper is a (closed-form) bound on the MSE performance of CM-minimizing estimates that is a simple function of the MSE performance of Weiner estimates. This bound, derived under the multi-source linear model in Fig. 1, provides the most formal link (established to date) between the CM and Wiener estimators, and as such, the most general testament to the MSE-robustness of the CM criterion.

II. Background

In this section, we give more detailed information on the linear system model and the MSE, UMSE, and CM criteria. The following notation is used throughout: $(\cdot)^t$

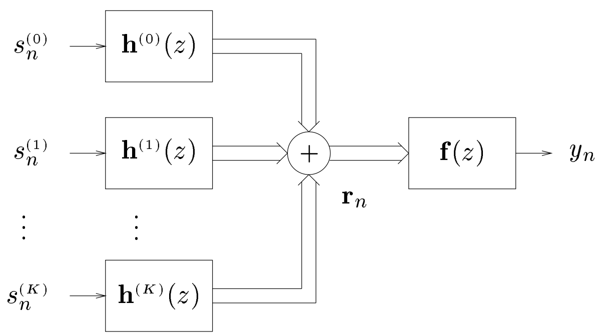

Fig. 1. Linear system model with K sources of interference.

denotes transpose, $(\cdot)^*$ conjugate, $(\cdot)^H$ hermitian, and $(\cdot)^\dagger$ Moore-Penrose pseudo-inverse. Likewise, $\mathrm{E}\{\cdot\}$ denotes expectation, $\|\mathbf{x}\|_p$ the p-norm defined by $\sqrt[p]{\sum_i |x_i|^p}$, $\mathbb{R}^+$ the field of non-negative real numbers, and $\ell_1(\mathbb{C})$ the space of absolutely-summable complex sequences. In general, we use boldface lowercase type to denote vector quantities and boldface uppercase type to denote matrix quantities.

A. Linear System Model

First we formalize the linear time-invariant multichannel model illustrated in Fig. 1. Say that the desired symbol sequence $\{s_n^{(0)}\}$ and K sources of interference $\{s_n^{(1)}\}, \dots, \{s_n^{(K)}\}$ each pass through separate linear "channels" before being observed at the receiver. In addition, say that the receiver uses a sequence of P-dimensional vector observations $\{\mathbf{r}_n\}$ to estimate (a possibly delayed version of) the desired source sequence, where the case $P > 1$ corresponds to a receiver that employs multiple sensors and/or samples at an integer multiple of the symbol rate. The observations $\mathbf{r}_n$ can be written $\mathbf{r}_n = \sum_{k=0}^{K} \sum_{i=0}^{\infty} \mathbf{h}_i^{(k)} s_{n-i}^{(k)}$, where $\{\mathbf{h}_n^{(k)}\}$ denote the impulse response coefficients of the linear time-invariant (LTI) channel $\mathbf{h}^{(k)}(z)$. We assume that $\mathbf{h}^{(k)}(z)$ is causal and bounded-input bounded-output (BIBO) stable with an auto-regressive moving-average (ARMA) structure.

From the vector-valued observation sequence $\{\mathbf{r}_n\}$, the receiver generates a sequence of linear estimates $\{y_n\}$ of $\{s_{n-\nu}^{(k)}\}$, where ν is a fixed integer. Using $\{\mathbf{f}_n\}$ to denote the impulse response of the linear estimator $\mathbf{f}(z)$, the estimates are formed as $y_n = \sum_{i=-\infty}^{\infty} \mathbf{f}_i^H \mathbf{r}_{n-i}$.We will assume that the linear system $\mathbf{f}(z)$ is BIBO stable with *constrained* ARMA structure.

In the sequel, we will focus almost exclusively on the combined channel-estimators $q^{(k)}(z) := \mathbf{f}^H(z)\mathbf{h}^{(k)}(z)$. The impulse response coefficients of $q^{(k)}(z)$ can be written

$$q_n^{(k)} \;=\; \sum_{i=-\infty}^{\infty} \mathbf{f}_i^H \mathbf{h}_{n-i}^{(k)}, \tag{3}$$

allowing the estimates to be written as $y_n = \sum_{k=0}^{K} \sum_{i=-\infty}^{\infty} q_i^{(k)} s_{n-i}^{(k)}$. Adopting the following vector notation helps to streamline the remainder of the paper.

$$\boldsymbol{q} := (\cdots, q_{-1}^{(0)}, \dots, q_{-1}^{(K)}, q_0^{(0)}, \dots, q_0^{(K)}, q_1^{(0)}, \dots, q_1^{(K)}, \cdots)^t,$$
$$\boldsymbol{s}(n) := (\cdots, s_{n+1}^{(0)}, \dots, s_{n+1}^{(K)}, s_n^{(0)}, \dots, s_n^{(K)}, s_{n-1}^{(0)}, \dots, s_{n-1}^{(K)}, \cdots)^t.$$

For instance, the estimates can be rewritten concisely as $y_n \;=\; \boldsymbol{q}^t \boldsymbol{s}(n)$.

It is important to recognize that that placing a particular structure on the channel and/or estimator will restrict the set of *attainable* channel-estimator responses, which we will denote by $\mathcal{Q}_{\mathrm{a}}$. For example, when the estimator is FIR, (3) implies that $\boldsymbol{q} \in \mathcal{Q}_{\mathrm{a}} = \mathrm{row}(\boldsymbol{\mathcal{H}})$, where

$$\boldsymbol{\mathcal{H}} := \begin{pmatrix} \mathbf{h}_0^{(0)} \cdots \mathbf{h}_0^{(K)} & \mathbf{h}_1^{(0)} \cdots \mathbf{h}_1^{(K)} & \mathbf{h}_2^{(0)} \cdots \mathbf{h}_2^{(K)} & \cdots \\ \mathbf{0} \;\cdots\; \mathbf{0} & \mathbf{h}_0^{(0)} \cdots \mathbf{h}_0^{(K)} & \mathbf{h}_1^{(0)} \cdots \mathbf{h}_1^{(K)} & \cdots \\ \vdots \qquad \vdots & \vdots \qquad \vdots & \vdots \qquad \vdots & \\ \mathbf{0} \;\cdots\; \mathbf{0} & \mathbf{0} \;\cdots\; \mathbf{0} & \mathbf{h}_0^{(0)} \cdots \mathbf{h}_0^{(K)} & \cdots \end{pmatrix}.$$

Restricting the estimator to be sparse or autoregressive, for example, would generate a different attainable set $\mathcal{Q}_{\mathrm{a}}$.

Throughout the paper, we make the following assumptions on the $K+1$ source processes:

S1) For all k, $\{s_n^{(k)}\}$ is zero-mean i.i.d.
S2) For $k \neq \ell$, $\{s_n^{(k)}\}$ is statistically independent of $\{s_n^{(\ell)}\}$.
S3) For all k, $\mathrm{E}\{|s_n^{(k)}|^2\} = \sigma_s^2$.
S4) $\mathcal{K}(s_n^{(0)}) < 0$, where $\mathcal{K}(\cdot)$ denotes kurtosis:

$$\mathcal{K}(s_n) \;:=\; \mathrm{E}\{|s_n|^4\} - 2\,\mathrm{E}^2\{|s_n|^2\} - \left|\mathrm{E}\{s_n^2\}\right|^2. \tag{4}$$

S5) If, for any k, $q^{(k)}(z)$ or $\{s_n^{(k)}\}$ is not real-valued, then $\mathrm{E}\{s_n^{(k)2}\} = 0$ for all k.

B. The Mean-Squared Error Criterion

The well-known mean-squared error (MSE) criterion was defined in (1) in terms of the estimate y_n and the estimand $s_{n-\nu}^{(0)}$. It is possible to derive closed-form expressions for the MMSE quantities that correspond our source and model assumptions [14]. In the case of an FIR estimator, for example, S1)–S3) imply $\boldsymbol{q}_{\mathrm{m},\nu} \;=\; \boldsymbol{\mathcal{H}}^t(\boldsymbol{\mathcal{H}}^*\boldsymbol{\mathcal{H}}^t)^\dagger \boldsymbol{\mathcal{H}}^* \mathbf{e}_\nu^{(0)}$, where $\mathbf{e}_\nu^{(k)}$ is a vector with single nonzero element of value 1 located so that $\boldsymbol{q}^t \mathbf{e}_\nu^{(k)} = q_\nu^{(k)}$. Henceforth, we denote MMSE quantities by the subscript "m."

C. Unbiased Mean-Squared Error

Since both symbol power and channel gain are unknown in the "blind" scenario, blind estimates suffer from a gain ambiguity. To ensure that our estimator performance evaluation is meaningful in the face of such ambiguity, we base our evaluation on normalized versions of the blind estimates and normalize by the receiver gain $q_\nu^{(0)}$. Given that the estimate y_n can be decomposed into signal and interference terms as

$$y_n = q_\nu^{(0)} s_{n-\nu}^{(0)} + \bar{\boldsymbol{q}}^t \bar{\boldsymbol{s}}(n), \tag{5}$$

where $\bar{\boldsymbol{q}}$ denotes $\boldsymbol{q}$ with the $q_\nu^{(0)}$ term removed and $\bar{\boldsymbol{s}}(n)$ denotes $\boldsymbol{s}(n)$ with the $s_{n-\nu}^{(0)}$ term removed, the normalized estimate $y_n/q_\nu^{(0)}$ can be referred to as "conditionally unbiased" since $\mathrm{E}\{y_n/q_\nu^{(0)}|s_{n-\nu}^{(0)}\} = s_{n-\nu}^{(0)}$.

The (conditionally) unbiased MSE (UMSE) associated with y_n, an estimate of $s_{n-\nu}^{(0)}$, is then defined

$$J_{\mathrm{u},\nu}(y_n) \ := \ \mathrm{E}\big\{|y_n/q_\nu^{(0)} - s_{n-\nu}^{(0)}|^2\big\}. \qquad (6)$$

Substituting (5) into (6), we find that

$$J_{\mathrm{u},\nu}(\boldsymbol{q}) \ = \ \mathrm{E}\big\{|\bar{\boldsymbol{q}}^t\bar{\boldsymbol{s}}(n)|^2\big\}/|q_\nu^{(0)}|^2 \ = \ \sigma_s^2\|\bar{\boldsymbol{q}}\|_2^2/|q_\nu^{(0)}|^2. \qquad (7)$$

Note that UMSE equals inverse signal to interference-plus-noise ratio (SINR), i.e., $J_{\mathrm{u},\nu} = \sigma_s^2\,\mathrm{SINR}_\nu^{-1}$ where

$$\mathrm{SINR}_\nu \ := \ \frac{\mathrm{E}\big\{|q_\nu^{(0)}s_{n-\nu}^{(0)}|^2\big\}}{\mathrm{E}\big\{|\bar{\boldsymbol{q}}^t\bar{\boldsymbol{s}}(n)|^2\big\}} \ = \ \frac{|q_\nu^{(0)}|^2}{\|\bar{\boldsymbol{q}}\|_2^2}.$$

D. The Constant Modulus Criterion

The constant modulus (CM) criterion, introduced independently in [1] and [2], was defined in (2) in terms of the estimates $\{y_n\}$. In (2), γ is a positive parameter known as the "dispersion constant." Though γ is often chosen according to the marginal statistics of the desired source process (when known), we will see that the UMSE performance of CM-minimizing estimates is insensitive to γ.

In the two "ideal" situations below, CM-minimizing estimates $\{y_n\}$ are known to take the form $y_n = \alpha s_{n-\nu}^{(0)}$, where $\alpha = e^{j\phi}\sqrt{\gamma\sigma_s^2/\,\mathrm{E}\{|s_n^{(0)}|^4\}}$, for some ϕ and ν. Note that these estimates have zero UMSE and, as such, are *perfect* up to a scalar ambiguity. For a single i.i.d. source that satisfies S4) and S5), this perfect CM-estimation property has been proven for

- unconstrained doubly-infinite estimators with BIBO channels [12], and
- causal FIR estimators with full-column rank $\boldsymbol{\mathcal{H}}$ [13].

In Section III-C, we extend the perfect CM-estimation property to the multi-source linear model described in Section II-A. See [15] for other properties of the CM criterion.

III. CM Performance under General Additive Interference

An algorithm for bounding the MSE performance of CM minimizers has been derived by Zeng et al. for the case of a real-valued i.i.d. source, a FIR channel, AWGN, and a finite-length estimator. The development for full column-rank (FCR) and non-FCR $\boldsymbol{\mathcal{H}}$ appear in [6] and [11], respectively. Though the relatively complicated Zeng algorithm generates reasonably tight CM-UMSE upper bounds, we have found that it is possible to derive tight bounds for the UMSE of CM-minimizing symbol estimates that

- have a closed-form expression,
- support arbitrary additive interference,
- support complex-valued channels and estimators, and
- support IIR (as well as FIR) channels and estimators.

We will now derive such bounds. Section III-A outlines our approach, Section III-B presents the main results, and Section III-C comments on these results.

A. The CM-UMSE Bounding Strategy

Say that $\boldsymbol{q}_{\mathrm{r},\nu}$ is an attainable reference channel-estimator response for the desired user ($k=0$) at delay ν. Formally, $\boldsymbol{q}_{\mathrm{r},\nu} \in \mathcal{Q}_\mathrm{a} \cap \mathcal{Q}_\nu^{(0)}$, where

$$\mathcal{Q}_\nu^{(0)} \ := \ \Big\{\boldsymbol{q} \text{ s.t. } |q_\nu^{(0)}| > \max_{(k,\delta)\neq(0,\nu)} |q_\delta^{(k)}|\Big\}.$$

$\mathcal{Q}_\nu^{(0)}$ defines the set of channel-estimators associated[1] with user 0 at delay ν. The set[2] of locally CM-minimizing channel-estimator responses for the desired user at delay ν will be denoted by $\{\boldsymbol{q}_{\mathrm{c},\nu}\}$ and defined as:

$$\{\boldsymbol{q}_{\mathrm{c},\nu}\} \ := \ \Big\{\arg\min_{\boldsymbol{q}\in\mathcal{Q}_\mathrm{a}} J_\mathrm{c}(\boldsymbol{q})\Big\} \cap \mathcal{Q}_\nu^{(0)}.$$

In general, it is not possible to determine closed-form expressions for $\{\boldsymbol{q}_{\mathrm{c},\nu}\}$, making it difficult to evaluate the UMSE of CM-minimizing estimates.

When $\boldsymbol{q}_{\mathrm{r},\nu}$ is in the vicinity of a $\boldsymbol{q}_{\mathrm{c},\nu}$ (the meaning of which will be made more precise later) then, by definition, this $\boldsymbol{q}_{\mathrm{c},\nu}$ must have CM cost less than or equal to the cost at $\boldsymbol{q}_{\mathrm{r},\nu}$. In this case, $\boldsymbol{q}_{\mathrm{c},\nu} \in \mathcal{Q}_\mathrm{c}(\boldsymbol{q}_{\mathrm{r},\nu})$, where

$$\mathcal{Q}_\mathrm{c}(\boldsymbol{q}_{\mathrm{r},\nu}) \ := \ \{\boldsymbol{q} \ \text{ s.t. } \ J_\mathrm{c}(\boldsymbol{q}) \leq J_\mathrm{c}(\boldsymbol{q}_{\mathrm{r},\nu})\} \cap \mathcal{Q}_\nu^{(0)}. \qquad (8)$$

This approach implies the following CM-UMSE upper bound:

$$J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{c},\nu}) \ \leq \ \max_{\boldsymbol{q}\in\mathcal{Q}_\mathrm{c}(\boldsymbol{q}_{\mathrm{r},\nu})} J_{\mathrm{u},\nu}(\boldsymbol{q}). \qquad (9)$$

Note that the maximization on the right of (9) does not explicitly involve the admissibility constraint $\mathcal{Q}_\mathrm{a}$; the constraint is implicitly incorporated through $\boldsymbol{q}_{\mathrm{r},\nu}$.

The tightness of the upper bound (9) will depend on the size and shape of $\mathcal{Q}_\mathrm{c}(\boldsymbol{q}_{\mathrm{r},\nu})$, motivating careful selection of the reference $\boldsymbol{q}_{\mathrm{r},\nu}$. Notice that the size of $\mathcal{Q}_\mathrm{c}(\boldsymbol{q}_{\mathrm{r},\nu})$ can usually be reduced via replacement of $\boldsymbol{q}_{\mathrm{r},\nu}$ with $\beta_\mathrm{r}\boldsymbol{q}_{\mathrm{r},\nu}$, where $\beta_\mathrm{r} := \arg\min_\beta J_\mathrm{c}(\beta\boldsymbol{q}_{\mathrm{r},\nu})$. (Since $\mathcal{Q}_\mathrm{a}$ is a linear subspace, $\boldsymbol{q}_{\mathrm{r},\nu} \in \mathcal{Q}_\mathrm{a}\cap\mathcal{Q}_\nu^{(0)} \Rightarrow \beta_\mathrm{r}\boldsymbol{q}_{\mathrm{r},\nu} \in \mathcal{Q}_\mathrm{a}\cap\mathcal{Q}_\nu^{(0)}$.) This implies that the direction (rather than the size) of $\boldsymbol{q}_{\mathrm{r},\nu}$ is important; the tightness of the CM-UMSE bound (9) will depend on collinearity of $\boldsymbol{q}_{\mathrm{r},\nu}$ and $\{\boldsymbol{q}_{\mathrm{c},\nu}\}$. Fig. 2 presents an illustration of this idea.

[1] Note that under S1)–S3), a particular {user, delay} combination is "associated" with an estimate if and only if that {user, delay} contributes more energy to the estimate than any other {user, delay}.

[2] We refer to the CM-minimizing channel-estimators in plural to avoid establishing the uniqueness of CM local minima within $\mathcal{Q}_\mathrm{a} \cap \mathcal{Q}_\nu^{(0)}$.

Zeng [11] has shown that in the case of an i.i.d. source, a FIR channel and AWGN noise, $\boldsymbol{q}_{\mathrm{c},\nu}$ are nearly collinear to the MMSE channel-estimator response $\boldsymbol{q}_{\mathrm{m},\nu}$. These findings, together with the abundant interpretations of the MMSE estimator and the existence of closed-form expressions for $\boldsymbol{q}_{\mathrm{m},\nu}$ suggest the reference choice $\boldsymbol{q}_{\mathrm{r},\nu} = \boldsymbol{q}_{\mathrm{m},\nu}$.

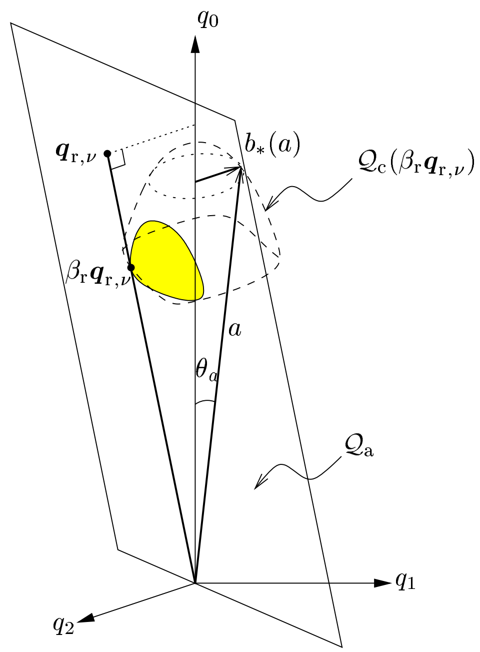

Fig. 2. Illustration of CM-UMSE upper-bounding technique using reference $\boldsymbol{q}_{\mathrm{r},\nu}$.

Determining a CM-UMSE upper bound from (9) can be accomplished as follows. Since both $J_{\mathrm{c}}(\boldsymbol{q})$ and $J_{\mathrm{u},\nu}(\boldsymbol{q})$ are invariant to phase rotation of $\boldsymbol{q}$ (i.e., scalar multiplication of $\boldsymbol{q}$ by $e^{j\phi}$ for $\phi \in \mathbb{R}$), we can restrict our attention to the set of "de-rotated" channel-estimator responses $\{\boldsymbol{q} \text{ s.t. } q_\nu^{(0)} \in \mathbb{R}^+\}$. Such $\boldsymbol{q}$ allow parameterization in terms of gain $a = \|\boldsymbol{q}\|_2$ and interference response $\bar{\boldsymbol{q}}$ (defined in Section II-C), where $\|\bar{\boldsymbol{q}}\|_2 \le a$. In terms of the pair $(a, \bar{\boldsymbol{q}})$, the upper bound in (9) may then be rewritten

$$\max_{\boldsymbol{q} \in \mathcal{Q}_{\mathrm{c}}(\beta_{\mathrm{r}}\boldsymbol{q}_{\mathrm{r},\nu})} J_{\mathrm{u},\nu}(\boldsymbol{q}) \;=\; \max_a \left(\max_{\bar{\boldsymbol{q}}:\,(a,\bar{\boldsymbol{q}}) \in \mathcal{Q}_{\mathrm{c}}(\beta_{\mathrm{r}}\boldsymbol{q}_{\mathrm{r},\nu})} J_{\mathrm{u},\nu}(a, \bar{\boldsymbol{q}}) \right).$$

Under particular conditions on the gain a and the reference $\boldsymbol{q}_{\mathrm{r},\nu}$ (made explicit in Section III-B), there exists a minimum interference gain

$$b_*(a) \;:=\; \min_{b(a)} \text{ s.t. } \Big\{ (a, \bar{\boldsymbol{q}}) \in \mathcal{Q}_{\mathrm{c}}(\beta_{\mathrm{r}}\boldsymbol{q}_{\mathrm{r},\nu}) \;\Rightarrow\; \|\bar{\boldsymbol{q}}\|_2 \le b(a) \Big\}, \tag{10}$$

which can be used in the containment:

$$\Big\{ (a, \bar{\boldsymbol{q}}) \in \mathcal{Q}_{\mathrm{c}}(\beta_{\mathrm{r}}\boldsymbol{q}_{\mathrm{r},\nu}) \Big\} \;\subset\; \Big\{ (a, \bar{\boldsymbol{q}}) \;\text{ s.t. }\; \|\bar{\boldsymbol{q}}\|_2 \le b_*(a) \Big\},$$

implying

$$\max_{\bar{\boldsymbol{q}}:\,(a,\bar{\boldsymbol{q}}) \in \mathcal{Q}_{\mathrm{c}}(\beta_{\mathrm{r}}\boldsymbol{q}_{\mathrm{r},\nu})} J_{\mathrm{u},\nu}(a, \bar{\boldsymbol{q}}) \;\le\; \max_{\bar{\boldsymbol{q}}:\,\|\bar{\boldsymbol{q}}\|_2 \le b_*(a)} J_{\mathrm{u},\nu}(a, \bar{\boldsymbol{q}}).$$

Applying (7) to the previous statement yields

$$\begin{aligned} \max_{\bar{\boldsymbol{q}}:\,\|\bar{\boldsymbol{q}}\|_2 \le b_*(a)} J_{\mathrm{u},\nu}(a, \bar{\boldsymbol{q}}) &= \max_{\bar{\boldsymbol{q}}:\,\|\bar{\boldsymbol{q}}\|_2 \le b_*(a)} \left(\frac{\|\bar{\boldsymbol{q}}\|_2^2}{a^2 - \|\bar{\boldsymbol{q}}\|_2^2} \right) \sigma_s^2 \\ &= \left(\frac{b_*^2(a)}{a^2 - b_*^2(a)} \right) \sigma_s^2, \end{aligned}$$

and putting these arguments together, we arrive at the CM-UMSE bound

$$J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{c},\nu}) \;\le\; \max_a \left(\frac{b_*^2(a)}{a^2 - b_*^2(a)} \right) \sigma_s^2. \tag{11}$$

The roles of various quantities can be summarized using Fig. 2. Starting with the arbitrarily-chosen (but attainable) reference channel-estimator response $\boldsymbol{q}_{\mathrm{r},\nu}$, the scalar β_{r} minimizes the CM cost that characterizes all scaled versions of $\boldsymbol{q}_{\mathrm{r},\nu}$. Since the CM minimum $\boldsymbol{q}_{\mathrm{c},\nu}$ is known to lie within the set $\mathcal{Q}_{\mathrm{c}}(\beta_{\mathrm{r}}\boldsymbol{q}_{\mathrm{r},\nu})$, delineated in Fig. 2 by long-dashed lines, the maximum UMSE within $\mathcal{Q}_{\mathrm{c}}(\beta_{\mathrm{r}}\boldsymbol{q}_{\mathrm{r},\nu})$ forms a valid upper bound for CM-UMSE.[3] Determining the maximum UMSE within $\mathcal{Q}_{\mathrm{c}}(\beta_{\mathrm{r}}\boldsymbol{q}_{\mathrm{r},\nu})$ is accomplished by first deriving $b_*(a)$, the smallest upper bound on interference gain for all $\boldsymbol{q} \in \mathcal{Q}_{\mathrm{c}}(\beta_{\mathrm{r}}\boldsymbol{q}_{\mathrm{r},\nu})$ that have a total gain of a, and then finding the particular combination of $\{a, b_*(a)\}$ that maximizes UMSE. The angle θ_a shown in Fig. 2 gives a simple trigonometric interpretation of the UMSE bound (11): $J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{c},\nu}) \le \max_a \tan^2(\theta_a)$. Also apparent from Fig. 2 is the notion that the valid range for a will depend on the choice of $\boldsymbol{q}_{\mathrm{r},\nu}$.

B. *Closed-Form Bounding Expressions for CM-UMSE*

After making a few definitions, we present CM-UMSE bounding expressions resulting from the approach of Section III-A. See [14] for detailed derivation and proofs.

Normalized kurtosis (not to be confused with $\mathcal{K}(\cdot)$ in (4)) is defined:

$$\kappa_s^{(k)} \;:=\; \mathrm{E}\big\{|s_n^{(k)}|^4\big\} / \,\mathrm{E}^2\big\{|s_n^{(k)}|^2\big\}. \tag{12}$$

Under the following definition of κ_g, our results will hold for both real-valued and complex-valued models.

$$\kappa_g \;:=\; \begin{cases} 3, & s_n^{(k)} \in \mathbb{R},\ \forall k, n \\ 2, & \text{otherwise}, \end{cases} \tag{13}$$

Note that, under S1) and S5), κ_g represents the normalized kurtosis of a Gaussian source. As shown in [14], the normalized and un-normalized kurtoses are related through

[3] Though a tighter bound would follow from use of the fact that $\exists\ \boldsymbol{q}_{\mathrm{c},\nu} \in \mathcal{Q}_{\mathrm{c}}(\beta_{\mathrm{r}}\boldsymbol{q}_{\mathrm{r},\nu}) \cap \mathcal{Q}_{\mathrm{a}}$ (denoted by the shaded area in Fig. 2), the set $\mathcal{Q}_{\mathrm{c}}(\beta_{\mathrm{r}}\boldsymbol{q}_{\mathrm{r},\nu}) \cap \mathcal{Q}_{\mathrm{a}}$ is too difficult to describe analytically.

$\mathcal{K}(s_n^{(k)}) = (\kappa_s^{(k)} - \kappa_g)\sigma_s^4$ when S3) and S5) hold. The following quantities are used in the sequel:

$$\kappa_s^{\min} := \min_{0\le k\le K} \kappa_s^{(k)}, \qquad \kappa_s^{\max} := \max_{0\le k\le K} \kappa_s^{(k)},$$
$$\rho_{\min} := \frac{\kappa_g - \kappa_s^{\min}}{\kappa_g - \kappa_s^{(0)}}, \qquad \rho_{\max} := \frac{\kappa_g - \kappa_s^{\max}}{\kappa_g - \kappa_s^{(0)}}.$$

Theorem 1: When there exists a Wiener estimator associated with the desired user at delay ν generating estimates with kurtosis $\kappa_{y_{\rm m}}$ obeying

$$\frac{\rho_{\min}}{1+\rho_{\min}} < \frac{\kappa_g - \kappa_{y_{\rm m}}}{\kappa_g - \kappa_s^{(0)}} \le 1,$$

there exists a CM-minimizing estimator that generates estimates associated with the same user/delay whose UMSE can be upper bounded by $J_{{\rm u},\nu}\big|_{{\rm c},\nu}^{\max,\kappa_{y_{\rm m}}}$, where

$$J_{{\rm u},\nu}\Big|_{{\rm c},\nu}^{\max,\kappa_{y_{\rm m}}} := \frac{1-\sqrt{(\rho_{\min}+1)\frac{\kappa_g-\kappa_{y_{\rm m}}}{\kappa_g-\kappa_s^{(0)}}-\rho_{\min}}}{\rho_{\min}+\sqrt{(\rho_{\min}+1)\frac{\kappa_g-\kappa_{y_{\rm m}}}{\kappa_g-\kappa_s^{(0)}}-\rho_{\min}}}\,\sigma_s^2. \quad (14)$$

While Theorem 1 presents a closed-form CM-UMSE bounding expression in terms of the kurtosis of the MMSE estimates, it is also possible to derive lower and upper bounds in terms of the UMSE of MMSE estimates.

Theorem 2: If Wiener UMSE $J_{{\rm u},\nu}(\boldsymbol{q}_{{\rm m},\nu}) < J_{\rm o}\sigma_s^2$, where

$$J_{\rm o} := \begin{cases} 2\sqrt{(1+\rho_{\min})^{-1}} - 1 & \kappa_s^{\max} \le \kappa_g \\ \frac{1-\sqrt{1-(3-\rho_{\min})(1+\rho_{\max})/4}}{\rho_{\max}+\sqrt{1-(3-\rho_{\min})(1+\rho_{\max})/4}}, & \kappa_s^{\max} > \kappa_g,\ \rho_{\max} \ne -1, \\ \frac{3-\rho_{\min}}{5+\rho_{\min}} & \kappa_s^{\max} > \kappa_g,\ \rho_{\max} = -1. \end{cases} \quad (15)$$

there exists a CM-minimizing estimator associated with the desired user at delay ν whose UMSE can be upper bounded by $J_{{\rm u},\nu}\big|_{{\rm c},\nu}^{\max,J_{{\rm u},\nu}(\boldsymbol{q}_{{\rm m},\nu})}$, where

$$J_{{\rm u},\nu}\Big|_{{\rm c},\nu}^{\max,J_{{\rm u},\nu}(\boldsymbol{q}_{{\rm m},\nu})} := \quad (16)$$
$$\begin{cases} \dfrac{1-\sqrt{(1+\rho_{\min})\left(1+\frac{J_{{\rm u},\nu}(\boldsymbol{q}_{{\rm m},\nu})}{\sigma_s^2}\right)^{-2}-\rho_{\min}}}{\rho_{\min}+\sqrt{(1+\rho_{\min})\left(1+\frac{J_{{\rm u},\nu}(\boldsymbol{q}_{{\rm m},\nu})}{\sigma_s^2}\right)^{-2}-\rho_{\min}}}\sigma_s^2 \\ \hfill \text{when } \kappa_s^{\max} \le \kappa_g, \\ \dfrac{1-\sqrt{(1+\rho_{\min})\left(1+\frac{J_{{\rm u},\nu}(\boldsymbol{q}_{{\rm m},\nu})}{\sigma_s^2}\right)^{-2}\left(1+\rho_{\max}\frac{J_{{\rm u},\nu}^2(\boldsymbol{q}_{{\rm m},\nu})}{\sigma_s^4}\right)-\rho_{\min}}}{\rho_{\min}+\sqrt{(1+\rho_{\min})\left(1+\frac{J_{{\rm u},\nu}(\boldsymbol{q}_{{\rm m},\nu})}{\sigma_s^2}\right)^{-2}\left(1+\rho_{\max}\frac{J_{{\rm u},\nu}^2(\boldsymbol{q}_{{\rm m},\nu})}{\sigma_s^4}\right)-\rho_{\min}}}\sigma_s^2 \\ \hfill \text{when } \kappa_s^{\max} > \kappa_g. \end{cases}$$

Note that the two cases of $J_{\rm o}$ in (15) and of $J_{{\rm u},\nu}\big|_{{\rm c},\nu}^{\max,J_{{\rm u},\nu}(\boldsymbol{q}_{{\rm m},\nu})}$ in (16) coincide as $\kappa_s^{\max} \to \kappa_g$. Equation (16) leads to an elegant approximation of the *extra* UMSE of CM-minimizing estimates:

$$\mathcal{E}_{{\rm u},\nu}(\boldsymbol{q}_{{\rm c},\nu}) := J_{{\rm u},\nu}(\boldsymbol{q}_{{\rm c},\nu}) - J_{{\rm u},\nu}(\boldsymbol{q}_{{\rm m},\nu}).$$

Theorem 3: If $J_{{\rm u},\nu}(\boldsymbol{q}_{{\rm m},\nu}) < J_{\rm o}\sigma_s^2$, then the extra UMSE of CM-minimizing estimates can be bounded as $\mathcal{E}_{{\rm u},\nu}(\boldsymbol{q}_{{\rm c},\nu}) \le \mathcal{E}_{{\rm u},\nu}\big|_{{\rm c},\nu}^{\max,J_{{\rm u},\nu}(\boldsymbol{q}_{{\rm m},\nu})}$, where

$$\begin{aligned} &\mathcal{E}_{{\rm u},\nu}\Big|_{{\rm c},\nu}^{\max,J_{{\rm u},\nu}(\boldsymbol{q}_{{\rm m},\nu})} \quad (17) \\ &:= J_{{\rm u},\nu}\Big|_{{\rm c},\nu}^{\max,J_{{\rm u},\nu}(\boldsymbol{q}_{{\rm m},\nu})} - J_{{\rm u},\nu}(\boldsymbol{q}_{{\rm m},\nu}) \\ &= \begin{cases} \frac{1}{2\sigma_s^2}\rho_{\min}J_{{\rm u},\nu}^2(\boldsymbol{q}_{{\rm m},\nu}) + \mathcal{O}\big(J_{{\rm u},\nu}^3(\boldsymbol{q}_{{\rm m},\nu})\big) \\ \hfill \text{when } \kappa_s^{\max} \le \kappa_g \\ \frac{1}{2\sigma_s^2}(\rho_{\min}-\rho_{\max})J_{{\rm u},\nu}^2(\boldsymbol{q}_{{\rm m},\nu}) + \mathcal{O}\big(J_{{\rm u},\nu}^3(\boldsymbol{q}_{{\rm m},\nu})\big) \\ \hfill \text{when } \kappa_s^{\max} > \kappa_g. \end{cases} \end{aligned}$$

Equation (17) implies that the extra UMSE of CM-minimizing estimates is upper bounded by approximately the *square* of the minimum UMSE. Fig. 3 plots the upper bound on CM-UMSE and extra CM-UMSE from (16) as a function of $J_{{\rm u},\nu}(\boldsymbol{q}_{{\rm m},\nu})/\sigma_s^2$ for various values of $\rho_{\min}$ and $\rho_{\max}$. The second-order approximation based on (17) appears very good for all but the largest values of UMSE.

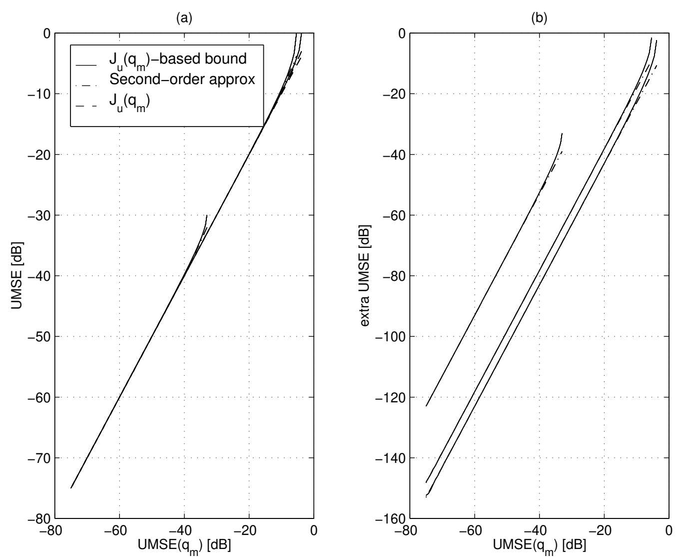

Fig. 3. Upper bound on (a) CM-UMSE and (b) extra CM-UMSE versus $J_{{\rm u},\nu}(\boldsymbol{q}_{{\rm m},\nu})$ (when $\sigma_s^2 = 1$) from (16) with second-order approximation from (17). From left to right, $\{\rho_{\min}, \rho_{\max}\} = \{1000, 0\}$, $\{1, -2\}$, and $\{1, 0\}$.

C. Comments on the CM-UMSE Bounds

C.1 Implicit Incorporation of $\mathcal{Q}_{\rm a}$

First, recall that the CM-UMSE bounding procedure incorporated $\mathcal{Q}_{\rm a}$, the set of attainable channel-estimators, *only* in the requirement that $\boldsymbol{q}_{{\rm r},\nu} \in \mathcal{Q}_{\rm a}$. Thus Theorems 1–3, written under the reference choice $\boldsymbol{q}_{{\rm r},\nu} = \boldsymbol{q}_{{\rm m},\nu} \in \mathcal{Q}_{\rm a} \cap \mathcal{Q}_\nu^{(0)}$, implicitly incorporate the channel and/or estimator constraints that define $\mathcal{Q}_{\rm a}$. For example, if $\boldsymbol{q}_{{\rm m},\nu}$

is the MMSE channel-estimator response constrained to the set of causal IIR estimators, then CM-UMSE bounds based on this $\boldsymbol{q}_{\mathrm{m},\nu}$ will implicitly incorporate the causality constraint. The implicit incorporation of attainable set $\mathcal{Q}_{\mathrm{a}}$ makes these bounding theorems general and easy to apply.

C.2 Effect of $\rho_{\min}$

When $\kappa_s^{\max} \leq \kappa_g$ and $\rho_{\min} = \frac{\kappa_g - \kappa_s^{\min}}{\kappa_g - \kappa_s^{(0)}} = 1$, the expressions in Theorems 1–3 simplify: e.g., $J_{\mathrm{o}} = (\sqrt{2}-1)$ and

$$
\begin{aligned}
J_{\mathrm{u},\nu}\big|_{\mathrm{c},\nu}^{\max,\boldsymbol{q}_{\mathrm{m},\nu}} &= \frac{1 - \sqrt{2\left(1 + \frac{J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})}{\sigma_s^2}\right)^{-2} - 1}}{1 + \sqrt{2\left(1 + \frac{J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})}{\sigma_s^2}\right)^{-2} - 1}} \sigma_s^2 \\
&\approx J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu}) + \frac{1}{2\sigma_s^2} J_{\mathrm{u},\nu}^2(\boldsymbol{q}_{\mathrm{m},\nu}).
\end{aligned}
$$

Typical scenarios leading to $\rho_{\min} = 1$ include

a) sub-Gaussian desired source in the presence of AWGN,
b) constant-modulus desired source in the presence of non-super-Gaussian interference, or
c) i.i.d. sources/interferers in the presence of AWGN.

The case $\rho_{\min} > 1$, on the other hand, might arise from the use of dense (and/or shaped) source constellations in the presence of interfering sources that are "more sub-Gaussian." In fact, source assumption S4) allows for arbitrarily large $\rho_{\min}$, which could result from a nearly-Gaussian desired source in the presence of non-Gaussian interference. Though Theorems 1–3 remain valid for arbitrarily high $\rho_{\min}$, the requirements placed on $\boldsymbol{q}_{\mathrm{m},\nu}$ via J_{o} become more stringent (recall Fig. 3).

C.3 Generalization of Perfect CM-Estimation Property

Finally, we note that the $J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})$-based CM-UMSE bound in Theorem 2 implies that the perfect CM-estimation property, proven under more restrictive conditions in [12] and [13], extends to the general multi-source linear model of Fig. 1:

Corollary 1: The CM-minimizing estimates are perfect (up to a scaling) in the presence of K interferers under S1)-S5) and

- *for BIBO IIR channels, unconstrained and doubly-infinite estimators; or*
- *for FIR channels, FCR channel matrices $\boldsymbol{\mathcal{H}}$; or*
- *for arbitrary channels, perfect (up to a scaling) Wiener estimates.*

Proof: Under S1)-S3) and the channel conditions above, $J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu}) = 0$. Under S1)-S5), Theorem 2 says that $J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu}) = 0 \;\Rightarrow\; J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{c},\nu}) = 0$. Hence, the estimates are perfect up to a (fixed) scale factor. ∎

IV. Conclusions

In this paper we have presented, for the general multi-source linear model of Fig. 1, two closed-form bounding expressions for the UMSE of CM-minimizing estimates and a generalization of the perfect CM-estimation property. This work confirms the longstanding conjecture (see, e.g., [1], [2], and [3]) that the MSE performance of the CM estimator is robust to *general* linear channels and *general* (multi-source) additive interference. As such, our results supersede previous work demonstrating the MSE-robustness of CM-minimizing estimates in special cases (e.g., when only AWGN is present, when the channel does not provide adequate diversity, or when the estimator has an insufficient number of adjustable parameters).

References

[1] D.N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Comm.*, vol. 28, no.11, pp. 1867-75, Nov. 1980.

[2] J.R. Treichler and B.G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoust., Speech, and Signal Processing*, vol. ASSP-31, no.2, pp. 459-72, Apr. 1983.

[3] C.R. Johnson, Jr., P. Schniter, et al., "Blind equalization using the constant modulus criterion: A review," *Proc. IEEE special issue on Blind System Identification and Estimation*, vol. 86, no. 10, pp. 1927-50, Oct. 1998.

[4] J.J. Shynk and R.P. Gooch, "The constant modulus array for cochannel signal copy and direction finding," *IEEE Trans. on Signal Processing*, vol. 44, no. 3, pp. 652-60, Mar. 1996.

[5] I. Fijalkow, A. Touzni, and J.R. Treichler, "Fractionally spaced equalization using CMA: Robustness to channel noise and lack of disparity," *IEEE Trans. on Signal Processing*, vol. 45, no. 1, pp. 56-66, Jan. 1997.

[6] H.H. Zeng, L. Tong and C.R. Johnson, Jr., "Relationships between the constant modulus and Wiener receivers," *IEEE Trans. Inform. Theory*, vol. 44, no. 4, pp. 1523-38, July 1998.

[7] D. Liu and L. Tong, "An analysis of constant modulus algorithm for array signal processing," *Signal Processing*, vol. 73, pp. 81-104, 1999.

[8] Y. Li, KJR Liu, and Z. Ding, "Length and cost dependent local minima of unconstrained blind channel equalizers," *IEEE Trans. on Signal Processing*, vol. 44, no. 11, pp. 2726-35, Nov. 1996.

[9] T.J. Endres, B.D.O. Anderson, C.R. Johnson, Jr., and M. Green, "Robustness to fractionally-spaced equalizer length using the constant modulus criterion," *IEEE Trans. on Signal Processing*, vol. 47, no. 2, pp. 544-9, Feb. 1999.

[10] P. Regalia and M. Mboup, "Undermodeled equalization: A characterization of stationary points for a family of blind criteria," *IEEE Trans. on Signal Processing*, vol. 47, no. 3, pp. 760-70, Mar. 1999.

[11] H.H. Zeng, L. Tong, and C.R. Johnson, Jr., "An analysis of constant modulus receivers," *IEEE Trans. on Signal Processing*, vol. 47, no. 11, pp. 2990-9, Nov. 1999.

[12] G.J. Foschini, "Equalizing without altering or detecting data (digital radio systems)," *AT&T Technical Journal*, vol. 64, no. 8, pp. 1885-911, Oct. 1985.

[13] Y. Li and Z. Ding, "Global convergence of fractionally spaced Godard (CMA) adaptive equalizers," *IEEE Trans. on Signal Processing*, vol. 44, no.4, pp. 818-26, Apr. 1996.

[14] P. Schniter and C.R. Johnson, Jr., "Bounds for the MSE performance of constant modulus estimators," to appear in *IEEE Trans. Inform. Theory*, 2000.

[15] C.R. Johnson, Jr., P. Schniter, I. Fijalkow, L. Tong, et al., and J. Treichler, "The core of FSE-CMA behavior theory," to appear in *Unsupervised Adaptive Filtering, Vol. 2: Blind Deconvolution*, (ed. Simon Haykin), New York, NY: Wiley, 2000.