Bounds for the MSE Performance of Constant Modulus Estimators

Philip Schniter and C. Richard Johnson, Jr., *Fellow, IEEE*

***Abstract*—The constant modulus (CM) criterion has become popular in the design of blind linear estimators of sub-Gaussian independent and identically distributed (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 estimators that depends only on the source kurtoses and the UMSE of Wiener estimators. Further analysis reveals that the extra UMSE of CM estimators can be upper-bounded by approximately the square of the Wiener (i.e., minimum) UMSE. Since our results hold for vector-valued finite-impulse response/infinite-impulse response (FIR/IIR) linear channels, vector-valued FIR/IIR estimators with a possibly constrained number of adjustable parameters, and multiple interferers with arbitrary distribution, they confirm the longstanding conjecture regarding the general mean-square error (MSE) robustness of CM estimators.**

***Index Terms*—Blind beamforming, blind deconvolution, blind equalization, blind multiuser detection, constant modulus algorithm, Godard algorithm.**

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 $\{\boldsymbol{h}^{(0)}(z), \ldots, \boldsymbol{h}^{(K)}(z)\}$. Our goal is to estimate the desired source using the (vector) linear estimator $\boldsymbol{f}(z)$. The linear estimates $\{y_n\}$ which minimize the mean-squared error (MSE)

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

are generated by the minimum MSE (MMSE) estimator, or Wiener estimator $\boldsymbol{f}_{m,\nu}(z)$. Specification of $\boldsymbol{f}_{m,\nu}(z)$, however, requires knowledge of the joint statistics of the observed sequence $\{\boldsymbol{r}_n\}$ 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 $\{\boldsymbol{r}_n\}$ are known, it may still be possible to estimate $\{s_n^{(0)}\}$ up to unknown magnitude and delay, i.e.,

$$y_n = \sum_i \boldsymbol{f}_i^H \boldsymbol{r}_{n-i} \approx \alpha = s_{n-\nu}^{(0)}$$

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). In [1], Shalvi and Weinstein showed that, when the sources and channels satisfy particular properties, perfect blind estimation is possible with knowledge of only the second- and fourth-order moments of the estimates $\{y_n\}$. Based on this observation, Shalvi and Weinstein proposed a blind estimation criterion based on minimizing the kurtosis of the estimates. It was later shown by Li and Ding [2] (and more recently by Regalia [3]) that the Shalvi–Weinstein (SW) estimators are closely related to the popular constant modulus (CM) estimators, proposed a decade earlier by Godard [4] and by Treichler and Larimore [5].

Minimization of the CM criterion has become perhaps the most studied and implemented means of blind equalization for data communication over dispersive channels (see, e.g., [6] and the references therein) and has also been used successfully as a means of blind beamforming (see, e.g., [7]). The CM criterion is defined below in terms of the estimates $\{y_n\}$ and a design parameter γ.

$$J_c(y_n) := E\{(|y_n|^2 - \gamma)^2\}. \tag{2}$$

The popularity of the CM criterion is usually attributed to

1) the existence of a simple adaptive algorithm ("CMA" [4], [5]) for estimation and tracking of the CM-minimizing estimator $\boldsymbol{f}_c(z)$, and
2) the excellent MSE performance of CM-minimizing estimators.

The second of these two points was first conjectured in the original works [4], [5] and provides the theme for the recently published comprehensive survey [6]. In this paper, we attempt to precisely quantify the general MSE performance of CM-minimizing estimators.

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) estimators 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., [8]–[10]);
2) channels that do not provide adequate diversity (e.g., [8], [11]); or
3) estimators with an insufficient number of adjustable parameters (e.g., [12], [13]).

A notable exception is the work of Zeng *et al.* [14], in which an algorithm is given to bound the MSE of CM-minimizing estimators 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 [14] is general

Manuscript received June 1, 1999; revised November 1, 1999. This work was supported in part by NSF under Grants MIP-9509011 and ECS-9811297, an Intel Foundation fellowship, and by Applied Signal Technology, Inc.

P. Schniter is with the Department of Electrical Engineering, The Ohio State University, Columbus OH 43210 USA (e-mail: schniter@ee.eng.ohio-state.edu).

C. R. Johnson is with Department of Electrical Engineering, Cornell University, Ithaca, NY 14853 USA (e-mail: johnson@ee.cornell.edu).

Communicated by U. Madhow, Associate Editor for Detection and Estimation.

Publisher Item Identifier S 0018-9448(00)09666-8.

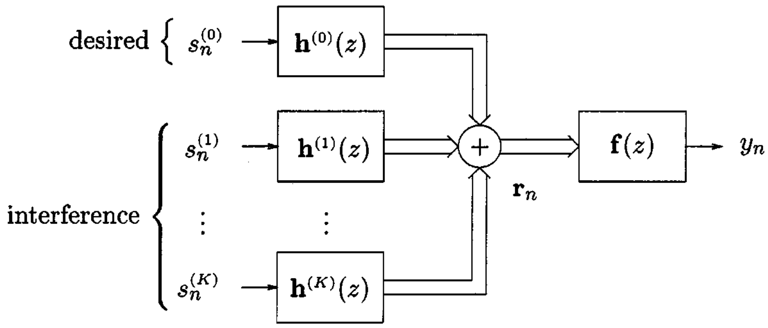


Fig. 1. Linear system model with K sources of interference.

enough to incorporate most combinations of the three conditions above, though not as general as the multisource model of Fig. 1. The bounding algorithm in [14] is rather involved, however, preventing a direct link between the MSE performance of CM and Wiener estimators (see Table I).

The main contribution of this paper is a (closed-form) bound on the MSE performance of CM-minimizing estimators that is a simple function of the MSE performance of Weiner estimators. This bound, derived under the multisource 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.

The organization of the paper is as follows. Section II discusses the properties of the system model and of the MSE and CM estimation criteria in detail, Section III derives the bound for the MSE performance of the CM criterion, and Section IV presents the results of numerical simulations demonstrating the efficacy of our bounding technique. Section V concludes the paper.

II. Background

In this section, we give more detailed information on the linear system model and the MSE, unbiased MSE, and CM criteria. The following notation is used throughout: $(\cdot)^t$ denotes transpose, $(\cdot)^*$ conjugate, $(\cdot)^H$ Hermitian, and $(\cdot)^\dagger$ Moore-Penrose pseudo-inverse. Likewise, $E\{\cdot\}$ denotes expectation, $\|\boldsymbol{x}\|_p$ the p-norm defined by $\sqrt[p]{\sum_i = |x_i|^p}$, $\|\boldsymbol{x}\|_{\boldsymbol{A}}$ the 2-norm defined by $\sqrt{\boldsymbol{x}^H \boldsymbol{A} x}$ for positive semidefinite $\boldsymbol{A}$, $\boldsymbol{I}$ the identity matrix, $\mathcal{Z}$ the z-transform, and $\mathbb{R}^+$ the field of nonnegative real numbers. 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 $\{s_n^{(1)}\}, \ldots, \{s_n^{(K)}\}$ each pass through separate linear "channels" before being observed at the receiver. The interference processes may correspond, e.g., to interference signals or additive noise processes. In addition, say that the receiver uses a sequence of P-dimensional vector observations $\{\boldsymbol{r}_n\}$ 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 $\boldsymbol{r}_n$ can be written

$$\boldsymbol{r}_n = \sum_{k=0}^{K} \sum_{i=0}^{\infty} \boldsymbol{h}_i^{(k)} s_{n-i}^{(k)} \tag{3}$$

where $\{\boldsymbol{h}_n^{(k)}\}$ denote the impulse response coefficients of the linear time-invariant (LTI) channel $\boldsymbol{h}^{(k)}(z)$. We assume that $\boldsymbol{h}^{(k)}(z)$ is causal and bounded-input bounded-output (BIBO) stable. Note that such $\boldsymbol{h}^{(k)}(z)$ admit infinite-impulse response (IIR) channel models.

From the vector-valued observation sequence $\{\boldsymbol{r}_n\}$, the receiver generates a sequence of linear estimates $\{y_n\}$ of $\{s_{n-\nu}^{(0)}\}$, where ν is a fixed integer. Using $\{\boldsymbol{f}_n\}$ to denote the impulse response of the linear estimator $\boldsymbol{f}(z)$, the estimates are formed as

$$y_n = \sum_{i=-\infty}^{\infty} \boldsymbol{f}_i^H \boldsymbol{r}_{n-i}. \tag{4}$$

We will assume that the linear system $\boldsymbol{f}(z)$ is BIBO stable with *constrained* autoregressive moving average (ARMA) structure, i.e., the pth element of $\boldsymbol{f}(z)$ takes the form

$$[\boldsymbol{f}(z)]_p = \frac{\sum_{i=0}^{L_b^{\langle p \rangle}} b_i^{\langle p \rangle} z^{-n_i^{\langle p \rangle}}}{1 + \sum_{i=1}^{L_a^{\langle p \rangle}} a_i^{\langle p \rangle} z^{-m_i^{\langle p \rangle}}}$$

where the $L_b^{\langle p \rangle} + 1$ "active" numerator coefficients $\{b_i^{\langle p \rangle}\}$, and the $L_a^{\langle p \rangle}$ active denominator coefficients $\{a_i^{\langle p \rangle}\}$ are constrained to the polynomial indexes $\{n_i^{\langle p \rangle}\}$ and $\{m_i^{\langle p \rangle}\}$, respectively.

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

$$q_n^{(k)} = \sum_{i=-\infty}^{\infty} \boldsymbol{f}_i^H \boldsymbol{h}_{n-i}^{(k)} \tag{5}$$

allowing the estimates to be written as

$$y_n = \sum_{k=0}^{K} \sum_{i=-\infty}^{\infty} q_i^{(k)} s_{n-i}^{(k)}. \tag{6}$$

TABLE I
ZENG ET AL.'S CM-UMSE BOUNDING ALGORITHM [14]

Assumptions:	The desired ($k=0$) channel is FIR with coefficients $\{\mathbf{h}_i^{(0)}\} \in \mathbb{R}^P$. AWGN of variance σ_w^2 is present at each of P sensors, so that $\mathbf{r}_n = \sum_{i=0}^{N_h-1} \left(\mathbf{h}_i^{(0)} s_{n-i}^{(0)} + \frac{\sigma_w}{\sigma_s} \sum_{k=1}^{P} \mathbf{e}_k s_{n-i}^{(k)} \right)$. The sources are real-valued and satisfy (S1)-(S5). The dispersion constant is $\gamma = \mathrm{E}\{\lvert s_n^{(0)}\rvert^4\}/\sigma_s^2$. The estimator $\boldsymbol{f} = (\mathbf{f}_0^t, \ldots, \mathbf{f}_{N_f-1}^t)^t$ has N_f coefficients of size $P \times 1$.
Definitions:	$\mathbf{H} := \begin{pmatrix} \mathbf{h}_0^{(0)} & \mathbf{h}_1^{(0)} & \cdots & \mathbf{h}_{N_h-1}^{(0)} & & & \\ & \ddots & & & \ddots & & \\ & & & \mathbf{h}_0^{(0)} & \mathbf{h}_1^{(0)} & \cdots & \mathbf{h}_{N_h-1}^{(0)} \end{pmatrix} \in \mathbb{R}^{PN_f \times (N_f+N_h-1)}$. $\mathbf{R} := \mathbf{H}\mathbf{H}^t + (\frac{\sigma_w}{\sigma_s})^2 \mathbf{I}$. $\boldsymbol{\Phi} := \mathbf{I} + (\frac{\sigma_w}{\sigma_s})^2 \left(\mathbf{H}^t\mathbf{H}\right)^{\dagger}$. With $\boldsymbol{\Phi} := \begin{pmatrix} \mathbf{C}_{11} & \mathbf{b}_1 & \mathbf{C}_{12} \\ \mathbf{b}_1^t & a & \mathbf{b}_2^t \\ \mathbf{C}_{12} & \mathbf{b}_2 & \mathbf{C}_{22} \end{pmatrix}$, set $a := [\boldsymbol{\Phi}]_{\nu,\nu}$, $\mathbf{b} := \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix}$ and $\mathbf{C} := \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12} & \mathbf{C}_{22} \end{pmatrix}$.
Calculations:	$\boldsymbol{q}_{\mathrm{m},\nu}^{(0)} = \mathbf{H}^t\mathbf{R}^{-1}\mathbf{H}\mathbf{e}_\nu$. $\boldsymbol{q}_{\mathrm{mI}}^{(0)} = \boldsymbol{q}_{\mathrm{m},\nu}^{(0)}(0:\nu-1,\ \nu+1:N_q-1)/\boldsymbol{q}_{\mathrm{m},\nu}^{(0)}(\nu)$, using MATLAB notation. $\alpha_{\mathrm{r},\nu} = \sqrt{\dfrac{\gamma \lVert \boldsymbol{q}_{\mathrm{m},\nu}^{(0)} \rVert_{\boldsymbol{\Phi}}^2}{3\lVert \boldsymbol{q}_{\mathrm{m},\nu}^{(0)} \rVert_{\boldsymbol{\Phi}}^4 - (3-\gamma)\lVert \boldsymbol{q}_{\mathrm{m},\nu}^{(0)} \rVert_4^4}}$. $\boldsymbol{q}_{\mathrm{r},\nu}^{(0)} = \alpha_{\mathrm{r},\nu}\boldsymbol{q}_{\mathrm{m},\nu}^{(0)}$. $J_{\mathrm{c},\nu}(\boldsymbol{q}_{\mathrm{r},\nu}^{(0)}) = 3\lVert \boldsymbol{q}_{\mathrm{r},\nu}^{(0)} \rVert_{\boldsymbol{\Phi}}^4 - 2\gamma \lVert \boldsymbol{q}_{\mathrm{r},\nu}^{(0)} \rVert_{\boldsymbol{\Phi}}^2 - (3-\gamma)\lVert \boldsymbol{q}_{\mathrm{r},\nu}^{(0)} \rVert_4^4 + \gamma^2$. $\boldsymbol{q}_{\mathrm{oI}}^{(0)} = -\mathbf{C}^{-1}\mathbf{b}$. $\theta_{\mathrm{o}} = (a - \mathbf{b}^t\mathbf{C}^{-1}\mathbf{b})^{-1}$. $\delta_{\mathrm{o}} = \lVert \boldsymbol{q}_{\mathrm{mI}}^{(0)} - \boldsymbol{q}_{\mathrm{oI}}^{(0)} \rVert_{\mathbf{C}}$.
UMSE Bound:	For the quartic polynomial $D(\delta) = c_1^2(\delta) - 4c_2(\delta)c_0$, where $c_0 = \gamma^2 - J_{\mathrm{c},\nu}(\boldsymbol{q}_{\mathrm{r},\nu}^{(0)})$, $c_1(\delta) = -2\gamma(\delta^2 + \theta_{\mathrm{o}}^{-1})$, $c_2(\delta) = 3(\delta^2 + \theta_{\mathrm{o}}^{-1})^2 - (3-\gamma)\big(1 + (\delta + \lVert \boldsymbol{q}_{\mathrm{oI}}^{(0)} \rVert_4)^4\big)$, find $\{\delta_1 < \cdots < \delta_m\}$ = real-valued roots of $D(\delta)$, and set $\delta_\star = \min\{\delta_i \mid \delta_i > \delta_{\mathrm{o}}\}$. If $\delta_\star \neq \emptyset$, $D(\delta_{\mathrm{o}}) \geq 0$, and $c_2(\delta) > 0$ for all $\delta \in [\delta_{\mathrm{o}}, \delta_\star]$, then $\mathrm{UMSE}(\boldsymbol{q}_{\mathrm{c},\nu}^{(0)}) \leq \delta_\star^2 + \theta_o^{-1} - 1$, else unable to compute bound.

Adopting the following vector notation helps to streamline the remainder of the paper.

$$
\begin{aligned}
\boldsymbol{q}^{(k)} &:= \left(\ldots, q_{-1}^{(k)}, q_0^{(k)}, q_1^{(k)}, \ldots\right)^t \\
\boldsymbol{q} &:= \Big(\ldots, q_{-1}^{(0)}, q_{-1}^{(1)}, \ldots, q_{-1}^{(K)}, \\
&\qquad q_0^{(0)}, q_0^{(1)}, \ldots, q_0^{(K)}, q_1^{(0)}, q_1^{(1)}, \ldots, q_1^{(K)}, \ldots\Big)^t \\
\boldsymbol{s}^{(k)}(n) &:= \left(\ldots, s_{n+1}^{(k)}, s_n^{(k)}, s_{n-1}^{(k)}, \ldots\right)^t \\
\boldsymbol{s}(n) &:= \Big(\ldots, s_{n+1}^{(0)}, s_{n+1}^{(1)}, \ldots, s_{n+1}^{(K)}, s_n^{(0)}, s_n^{(1)}, \ldots, s_n^{(K)}, \\
&\qquad s_{n-1}^{(0)}, s_{n-1}^{(1)}, \ldots, s_{n-1}^{(K)}, \ldots\Big)^t.
\end{aligned}
$$

For instance, the estimates can be rewritten concisely as

$$y_n = \sum_{k=0}^{K} \boldsymbol{q}^{(k)t}\boldsymbol{s}^{(k)}(n) = \boldsymbol{q}^t\boldsymbol{s}(n). \tag{7}$$

The source-specific unit vector $\boldsymbol{e}_\nu^{(k)}$ will also prove convenient. $\boldsymbol{e}_\nu^{(k)}$ is a column vector with a single nonzero element of value 1 located such that

$$\boldsymbol{q}^t\boldsymbol{e}_\nu^{(k)} = q_\nu^{(k)}.$$

We now point out two important properties of $\boldsymbol{q}$. First, it is important to recognize that placing a particular structure on the channel and/or estimator will restrict the set of *attainable* global responses, which we will denote by $\mathcal{Q}_a$. For example, when the estimator is FIR, (5) implies that $\boldsymbol{q} \in \mathcal{Q}_a = \text{row}(\boldsymbol{\mathcal{H}})$, where we get (8) (see the bottom of this page). Restricting the estimator to be sparse or autoregressive, for example, would generate a different attainable set $\mathcal{Q}_a$. Next, BIBO-stable $\boldsymbol{f}(z)$ and $\boldsymbol{h}^{(k)}(z)$ imply BIBO-stable $q^{(k)}(z)$, so that $\|\boldsymbol{q}^{(k)}\|_p$ exists for all $p \geq 1$, and thus $\|\boldsymbol{q}\|_p$ does as well.

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

- S1) For all k, $\{s_n^{(k)}\}$ is zero-mean independent and identically distributed (i.i.d).
- S2) $\{s_n^{(0)}\}, \ldots, \{s_n^{(K)}\}$ are jointly independent.
- S3) For all k, $E\{|s_n^{(k)}|^2\} = \sigma_s^2$.
- S4) $\mathcal{K}(s_n^{(0)}) < 0$, where $\mathcal{K}(\cdot)$ denotes kurtosis

$$\mathcal{K}(s_n) := E\{|s_n|^4\} - 2E^2\{|s_n|^2\} - \left|E\left\{s_n^2\right\}\right|^2. \tag{9}$$

- S5) If, for any k, $q^{(k)}(z)$ or $\{s_n^{(k)}\}$ is not real-valued, then $E\{{s_n^{(k)}}^2\} = 0$ 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 $s_{n-\nu}^{(0)}$. Using S1)–S3), we may rewrite (1) in terms of the global response $\boldsymbol{q}$

$$J_{m,\nu}(\boldsymbol{q}) = \left\|\boldsymbol{q} - \boldsymbol{e}_\nu^{(0)}\right\|_2^2 \sigma_s^2. \tag{10}$$

Denoting MMSE quantities by the subscript "m," it can be shown [15] that in the unconstrained (noncausal) IIR case, S1)–S3) imply that the MMSE response is

$$q_{m,\nu}^{(l)}(z) = z^{-\nu}\boldsymbol{h}^{(0)H}\left(\frac{1}{z^*}\right)\left(\sum_k \boldsymbol{h}^{(k)}(z)\boldsymbol{h}^{(k)H}\left(\frac{1}{z^*}\right)\right)^{\dagger}\boldsymbol{h}^{(l)}(z), \quad \text{for } \ell = 0, \ldots, K \tag{11}$$

while in the FIR case, S1)–S3) imply

$$\boldsymbol{q}_{m,\nu} = \boldsymbol{\mathcal{H}}^t(\boldsymbol{\mathcal{H}}^*\boldsymbol{\mathcal{H}}^t)^\dagger \boldsymbol{\mathcal{H}}^* \boldsymbol{e}_\nu^{(0)}.$$

In this latter case, $\boldsymbol{q}_{m,\nu}$ is the projection of $\boldsymbol{e}_\nu^{(0)}$ onto the row space of $\boldsymbol{\mathcal{H}}^*$.

C. Unbiased Mean-Squared Error (UMSE)

Since both symbol power and channel gain are unknown in the "blind" scenario, blind estimators 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 estimators where the normalization factor is the receiver gain $q_\nu^{(0)}$. Given that the estimate y_n can be decomposed into signal and interference terms as

$$y_n = q_\nu^{(0)} s_{n-\nu}^{(0)} + \bar{\boldsymbol{q}}^t \bar{\boldsymbol{s}}(n), \tag{12}$$

where

$$\bar{\boldsymbol{q}} := \text{``}\boldsymbol{q} \text{ with the } q_\nu^{(0)} \text{ term removed''}$$
$$\bar{\boldsymbol{s}}(n) := \text{``}\boldsymbol{s}(n) \text{ with the } s_{n-\nu}^{(0)} \text{ term removed''}$$

the normalized estimate $y_n/q_\nu^{(0)}$ can be referred to as "conditionally unbiased" since $E\{y_n/q_\nu^{(0)} \mid s_{n-\nu}^{(0)}\} = s_{n-\nu}^{(0)}$.

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

$$J_{u,\nu}(y_n) := E\left\{\left|y_n/q_\nu^{(0)} - s_{n-\nu}^{(0)}\right|^2\right\}. \tag{13}$$

Substituting (12) into (13), we find that

$$J_{u,\nu}(\boldsymbol{q}) = \frac{E\{|\bar{\boldsymbol{q}}^t\bar{\boldsymbol{s}}(n)|^2\}}{\left|q_\nu^{(0)}\right|^2} = \frac{\|\bar{\boldsymbol{q}}\|_2^2}{\left|q_\nu^{(0)}\right|^2}\sigma_s^2 \tag{14}$$

where the second equality invokes assumptions S1)–S3).

Note that UMSE is related to signal-to-interference-plus-noise ratio (SINR) via $J_{u,\nu} = \sigma_s^2 \text{SINR}_\nu^{-1}$, where

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

D. The Constant Modulus (CM) Criterion

The constant modulus (CM) criterion, introduced independently in [4] and [5], was defined in (2) in terms of the statistics of $\{y_n\}$. 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 they are known), we will see that the UMSE performance of CM-minimizing estimators 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

$$\alpha = e^{j\phi}\sqrt{\gamma\sigma_s^2/E\{|s_n^{(0)}|^4\}}$$

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 [16], and

$$\boldsymbol{\mathcal{H}} := \begin{pmatrix} \boldsymbol{h}_0^{(0)} & \cdots & \boldsymbol{h}_0^{(K)} & \boldsymbol{h}_1^{(0)} & \cdots & \boldsymbol{h}_1^{(K)} & \boldsymbol{h}_2^{(0)} & \cdots & \boldsymbol{h}_2^{(K)} & \cdots \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{h}_0^{(0)} & \cdots & \boldsymbol{h}_0^{(K)} & \boldsymbol{h}_1^{(0)} & \cdots & \boldsymbol{h}_1^{(K)} & \cdots \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots & \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{h}_0^{(0)} & \cdots & \boldsymbol{h}_0^{(K)} & \cdots \end{pmatrix}. \tag{8}$$

- causal FIR estimators with full-column rank (FCR) $\mathcal{H}$ [17], [18].

In Section III-C, we extend the perfect CM-estimation property to the multisource linear model described in Section II-A.

For more comprehensive surveys on the CM criterion, see [6] and [19].

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, an FIR channel, AWGN, and a finite-length estimator. The development for FCR and non-FCR $\mathcal{H}$ appear in [9] and [14], respectively. Using the notation established in Section II-A, the algorithm of [14] is summarized in Table I. Though the relatively complicated Zeng algorithm generates reasonably tight CM-UMSE upper bounds (as we shall see in Section IV), we have found that it is possible to derive tight bounds for the UMSE of CM-minimizing symbol estimators 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. Proof details appear in the Appendix.

A. The CM-UMSE Bounding Strategy

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

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

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

$$\{\boldsymbol{q}_{c,\nu}\} := \{\arg\min_{\boldsymbol{q}\in\mathcal{Q}_a} J_c(\boldsymbol{q})\} \cap \mathcal{Q}_\nu^{(0)}.$$

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

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

$$\mathcal{Q}_c(\boldsymbol{q}_{r,\nu}) := \{\boldsymbol{q} \text{ s.t. } J_c(\boldsymbol{q}) \leq J_c(\boldsymbol{q}_{r,\nu})\} \cap \mathcal{Q}_\nu^{(0)}. \tag{15}$$

This approach implies the following CM-UMSE upper bound:

$$J_{u,\nu}(\boldsymbol{q}_{c,\nu}) \leq \max_{\boldsymbol{q}\in\mathcal{Q}_c(\boldsymbol{q}_{r,\nu})} J_{u,\nu}(\boldsymbol{q}). \tag{16}$$

[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 responses as a set to avoid establishing the existence or uniqueness of CM local minima within $\mathcal{Q}_a \cap \mathcal{Q}_\nu^{(0)}$ at this time.

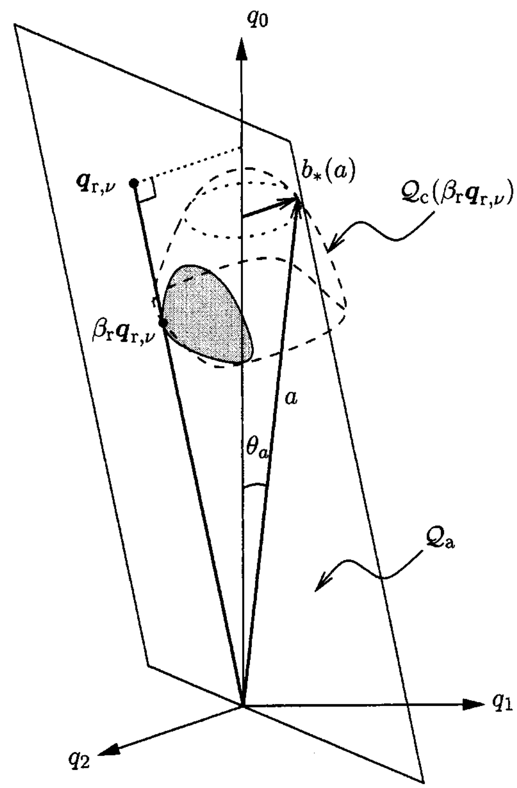


Fig. 2. Illustration of CM-UMSE upper-bounding technique using reference $\boldsymbol{q}_{r,\nu}$.

Note that the maximization on the right of (16) does not explicitly involve the attainibility constraint $\mathcal{Q}_a$; the constraint is implicitly incorporated through $\boldsymbol{q}_{r,\nu}$.

The tightness of the upper bound (16) will depend on the size and shape of $\mathcal{Q}_c(\boldsymbol{q}_{r,\nu})$, motivating careful selection of the reference $\boldsymbol{q}_{r,\nu}$. Notice that the size of $\mathcal{Q}_c(\boldsymbol{q}_{r,\nu})$ can usually be reduced via replacement of $\boldsymbol{q}_{r,\nu}$ with $\beta_r\boldsymbol{q}_{r,\nu}$, where $\beta_r := \arg\min_\beta J_c(\beta\boldsymbol{q}_{r,\nu})$. This implies that the direction (rather than the size) of $\boldsymbol{q}_{r,\nu}$ is important; the tightness of the CM-UMSE bound (16) will depend on collinearity of $\boldsymbol{q}_{r,\nu}$ and $\boldsymbol{q}_{c,\nu}$. Fig. 2 presents an illustration of this idea.

Zeng has shown that in the case of an i.i.d. source, an FIR channel, and AWGN noise, $\boldsymbol{q}_{c,\nu}$ is nearly collinear to the MMSE response $\boldsymbol{q}_{m,\nu}$ [14]. These findings, together with the abundant interpretations of the MMSE estimator and the existence of closed-form expressions for $\boldsymbol{q}_{m,\nu}$ suggest the reference choice $\boldsymbol{q}_{r,\nu} = \boldsymbol{q}_{m,\nu}$.

Determining a CM-UMSE upper bound from (16) can be accomplished as follows. Since both $J_c(\boldsymbol{q})$ and $J_{u,\nu}(\boldsymbol{q})$ 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" 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 $\|\bar{\boldsymbol{q}}\|_2 \leq a$. In terms of the pair $(a, \bar{\boldsymbol{q}})$, the upper bound in (16) may then be rewritten

$$\max_{\boldsymbol{q}\in\mathcal{Q}_c(\beta_r\boldsymbol{q}_{r,\nu})} J_{u,\nu}(\boldsymbol{q}) = \max_a \left(\max_{\bar{\boldsymbol{q}}:(a,\bar{\boldsymbol{q}})\in\mathcal{Q}_c(\beta_r\boldsymbol{q}_{r,\nu})} J_{u,\nu}(a,\bar{\boldsymbol{q}}) \right)$$

where

$$J_{u,\nu}(a,\bar{\boldsymbol{q}}) = \|\bar{\boldsymbol{q}}\|_2^2\sigma_s^2/(a^2 - \|\bar{\boldsymbol{q}}\|_2^2)$$

from (14). Under particular conditions on the gain a and the reference $\boldsymbol{q}_{r,\nu}$ (made explicit in Section III-B), there exists a minimum interference gain

$$b_*(a) := \min b(a) \text{ s.t. } \{(a,\bar{\boldsymbol{q}}) \in \mathcal{Q}_c(\beta_r \boldsymbol{q}_{r,\nu}) \Rightarrow \|\bar{\boldsymbol{q}}\|_2 \le b(a)\} \tag{17}$$

which can be used in the containment

$$\{(a,\bar{\boldsymbol{q}}) \in \mathcal{Q}_c(\beta_r \boldsymbol{q}_{r,\nu})\} \subset \{(a,\bar{\boldsymbol{q}}) \text{ s.t. } \|\bar{\boldsymbol{q}}\|_2 \le b_*(a)\}$$

implying

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

Applying (14) to the previous statement yields

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

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

$$J_{u,\nu}(\boldsymbol{q}_{c,\nu}) \le \max_a \left(\frac{b_*^2(a)}{a^2 - b_*^2(a)}\right)\sigma_s^2. \tag{18}$$

The roles of various quantities can be summarized using Fig. 2. Starting with the (attainable) global reference response $\boldsymbol{q}_{r,\nu}$, the scalar β_r minimizes the CM cost that characterizes all scaled versions of $\boldsymbol{q}_{r,\nu}$. Since the CM minimum $\boldsymbol{q}_{c,\nu}$ is known to lie within the set $\mathcal{Q}_c(\beta_r\boldsymbol{q}_{r,\nu})$, delineated in Fig. 2 by long-dashed lines, the maximum UMSE within $\mathcal{Q}_c(\beta_r\boldsymbol{q}_{r,\nu})$ forms a valid upper bound for CM-UMSE.[3] Determining the maximum UMSE within $\mathcal{Q}_c(\beta_r\boldsymbol{q}_{r,\nu})$ is accomplished by first deriving $b_*(a)$, the smallest upper bound on interference gain for all $\boldsymbol{q} \in \mathcal{Q}_c(\beta_r\boldsymbol{q}_{r,\nu})$ 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 (18): $J_{u,\nu}(\boldsymbol{q}_{c,\nu}) \le \max_a \tan^2(\theta_a)$. Also apparent from Fig. 2 is the notion that the valid range for a will depend on the choice of $\boldsymbol{q}_{r,\nu}$.

B. Derivation of the CM-UMSE Bounds

In this section, we derive CM-UMSE bounds based on the method described in Section III-A. The main steps in the derivation are presented as lemmas, with proofs appearing in the Appendix.

It is convenient to now define the *normalized* kurtosis (not to be confused with $\mathcal{K}(\cdot)$ in (9))

$$\kappa_s^{(k)} := \frac{E\left\{\left|s_n^{(k)}\right|^4\right\}}{E^2\left\{\left|s_n^{(k)}\right|^2\right\}}. \tag{19}$$

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

Under the following definition of κ_g:

$$\kappa_g := \begin{cases} 3, & s_n^{(k)} \in \mathbb{R},\ \forall k,n \\ 2, & \text{otherwise}\end{cases} \tag{20}$$

our results will hold for both real-valued and complex-valued models. Note that, under S1) and S5), κ_g represents the normalized kurtosis of a Gaussian source. As shown in Subsection A of the Appendix, the normalized and unnormalized kurtoses are related through $\mathcal{K}(s_n^{(k)}) = (\kappa_s^{(k)} - \kappa_g)\sigma_s^4$ when S3) and S5) hold. The following kurtosis-based quantities will all prove useful in the sequel:

$$\kappa_s^{\min} := \min_{0\le k\le K} \kappa_s^{(k)} \tag{21}$$

$$\kappa_s^{\max} := \max_{0\le k\le K} \kappa_s^{(k)} \tag{22}$$

$$\rho_{\min} := \frac{\kappa_g - \kappa_s^{\min}}{\kappa_g - \kappa_s^{(0)}} \tag{23}$$

$$\rho_{\max} := \frac{\kappa_g - \kappa_s^{\max}}{\kappa_g - \kappa_s^{(0)}}. \tag{24}$$

The first step is to express the CM cost (2) in terms of the global response $\boldsymbol{q}$ (defined in Section II-A).

Lemma 1: The CM cost may be written in terms of global response $\boldsymbol{q}$ as

$$\begin{aligned}\frac{J_c(\boldsymbol{q})}{\sigma_s^4} &= \sum_k \left(\kappa_s^{(k)} - \kappa_g\right)\|\boldsymbol{q}^{(k)}\|_4^4 \\ &\quad + \kappa_g\|\boldsymbol{q}\|_2^4 - 2\left(\gamma/\sigma_s^2\right)\|\boldsymbol{q}\|_2^2 + \left(\gamma/\sigma_s^2\right)^2.\end{aligned} \tag{25}$$

Similar expressions for the CM cost have been generated for the case of a desired user in AWGN (see, e.g., [6]).

The CM cost expression (25) can now be used to compute the CM cost at scaled versions of a reference $\boldsymbol{q}_{r,\nu}$.

Lemma 2: For any $\boldsymbol{q}_{r,\nu}$

$$\beta_r = \arg\min_\beta J_c(\beta\boldsymbol{q}_{r,\nu}) = \frac{1}{\|\boldsymbol{q}_{r,\nu}\|_2}\sqrt{\left(\frac{\gamma}{\sigma_s^2}\right)\frac{1}{\kappa_{y_r}}}$$

and

$$J_c(\beta_r\boldsymbol{q}_{r,\nu}) = \gamma^2\left(1 - \kappa_{y_r}^{-1}\right) \tag{26}$$

where κ_{y_r} is the normalized kurtosis of the estimates generated by the reference $\boldsymbol{q}_{r,\nu}$.

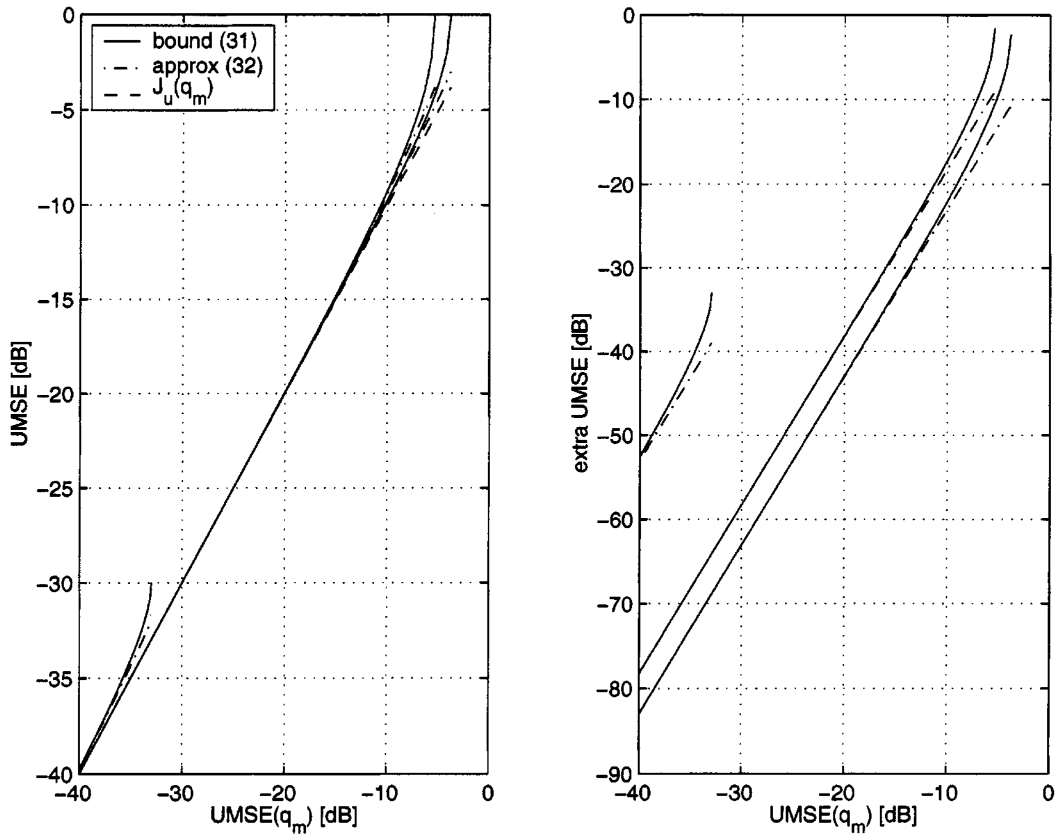
The expression for $J_c(\beta_r\boldsymbol{q}_{r,\nu})$ in (26) leads directly to an expression for $\mathcal{Q}_c(\beta_r\boldsymbol{q}_{r,\nu})$, from which the minimum interference gain $b_*(a)$ of (17) can be derived.

Lemma 3: The nonnegative gain $b_*(a)$ satisfying definition (17) can be upper-bounded as

$$b_*(a) \le a\sqrt{\frac{1 - \sqrt{1 - (\rho_{\min}+1)\frac{C(a,\boldsymbol{q}_{r,\nu})}{a^4}}}{\rho_{\min}+1}}, \quad \text{when } 0 \le \frac{C(a,\boldsymbol{q}_{r,\nu})}{a^4} \le \frac{3-\rho_{\min}}{4} \tag{27}$$

where $C(a,\boldsymbol{q}_{r,\nu})$ is defined in (42).

Equations (18) and (27) lead to an upper bound for the UMSE of CM-minimizing estimators.



(a) (b)

Fig. 3. Upper bound on (a) CM-UMSE and (b) extra CM-UMSE versus $J_{u,\nu}(\boldsymbol{q}_{m,\nu})$ (when $\sigma_s^2 = 1$) from (31) with second-order approximation from (32). From left to right, $\{\rho_{\min}, \rho_{\max}\} = \{1000, 0\}$, $\{1, -2\}$, and $\{1, 0\}$.

Theorem 1: When there exists a Wiener estimator associated with the desired user and delay ν generating estimates with kurtosis κ_{y_m} obeying

$$\frac{1+\rho_{\min}}{4} < \frac{\kappa_g - \kappa_{y_m}}{\kappa_g - \kappa_s^{(0)}} \le 1 \tag{28}$$

the UMSE of CM-minimizing estimators associated with the same user/delay can be upper-bounded by $J_{u,\nu}|_{c,\nu}^{\max,\kappa_{y_m}}$, where

$$J_{u,\nu}|_{c,\nu}^{\max,\kappa_{y_m}} := \frac{1 - \sqrt{(\rho_{\min}+1)\frac{\kappa_g - \kappa_{y_m}}{\kappa_g - \kappa_s^{(0)}} - \rho_{\min}}}{\rho_{\min} + \sqrt{(\rho_{\min}+1)\frac{\kappa_g - \kappa_{y_m}}{\kappa_g - \kappa_s^{(0)}} - \rho_{\min}}}\sigma_s^2. \tag{29}$$

Furthermore, (28) guarantees the existence of a CM-minimizing estimator associated with this user/delay when $\boldsymbol{q}$ is FIR.

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 the MMSE estimator.

Theorem 2: If Wiener UMSE $J_{u,\nu}(\boldsymbol{q}_{m,\nu}) < J_o\sigma_s^2$, where we get (30) (see the the bottom of the page). The UMSE of CM-minimizing estimators associated with the same user/delay can be upper-bounded as follows:

$$J_{u,\nu}(\boldsymbol{q}_{m,\nu}) \le J_{u,\nu}(\boldsymbol{q}_{c,\nu}) \le J_{u,\nu}|_{c,\nu}^{\max,\kappa_{y_m}} \le J_{u,\nu}|_{c,\nu}^{\max,J_{u,\nu}(\boldsymbol{q}_{m,\nu})}$$

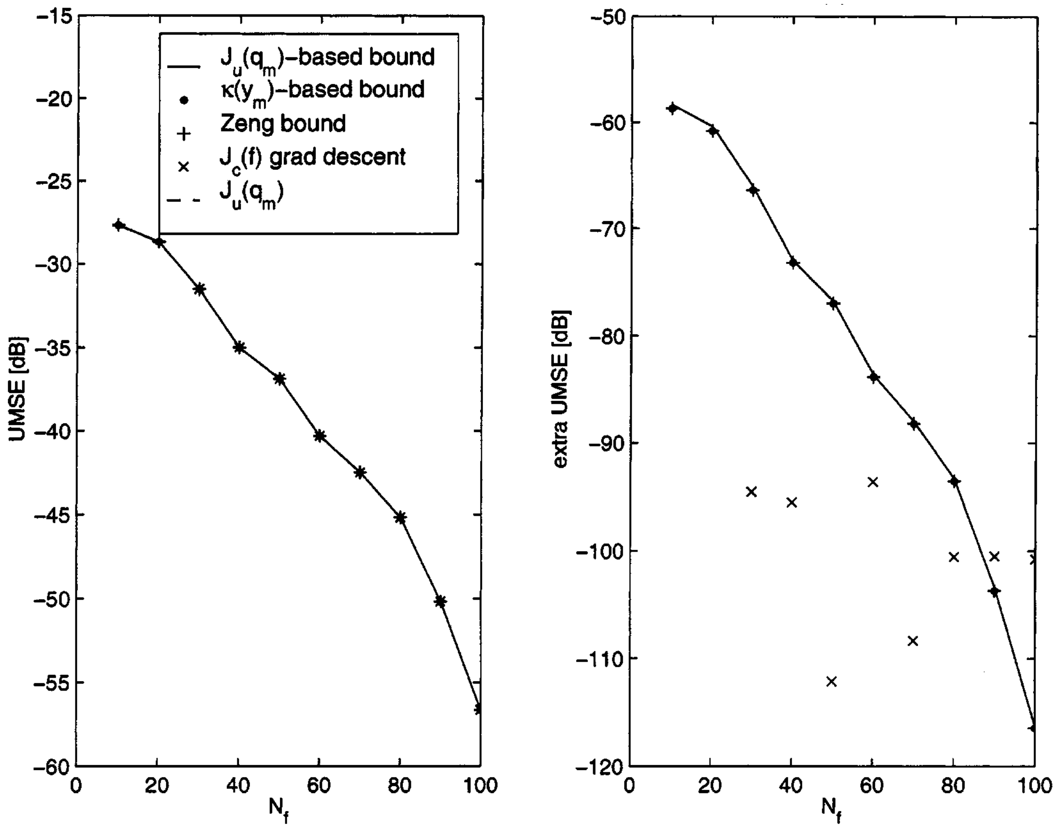
where we get (31) (see the bottom of the following page). Furthermore, (30) guarantees the existence of a CM-minimizing estimator associated with this user/delay when $\boldsymbol{q}$ is FIR.

Note that the two cases of J_o in (30) and of $J_{u,\nu}|_{c,\nu}^{\max,J_{u,\nu}(\boldsymbol{q}_{m,\nu})}$ in (31) coincide as $\kappa_s^{\max} \to \kappa_g$.

Equation (31) leads to an elegant approximation of the *extra* UMSE of CM-minimizing estimators

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

$$J_o := \begin{cases} 2\sqrt{(1+\rho_{\min})^{-1}} - 1 & \kappa_s^{\max} \le \kappa_g \\ \frac{1-\sqrt{1-(3-\rho_{\min})(1+\rho_{\max})/4}}{\rho_{\max}+\sqrt{1-(3-\rho_{\min})(1+\rho_{\max})/4}}, & \kappa_s^{\max} > \kappa_g,\ \rho_{\max} \ne -1 \\ \frac{3-\rho_{\min}}{5+\rho_{\min}}, & \kappa_s^{\max} > \kappa_g,\ \rho_{\max} = -1. \end{cases} \tag{30}$$



(a) (b)

Fig. 4. Bounds on CM-UMSE versus estimator length N_f for SPIB microwave channel #5 and 8-PAM.

Theorem 3: If $J_{u,\nu}(\boldsymbol{q}_{m,\nu}) < J_o \sigma_s^2$, then the extra UMSE of CM-minimizing estimators can be bounded as

$$\mathcal{E}_{u,\nu}(\boldsymbol{q}_{c,\nu}) \le \mathcal{E}_{u,\nu}\big|_{c,\nu}^{\max, J_{u,\nu}(\boldsymbol{q}_{m,\nu})}$$

where we get (32) at the bottom of this page.

Equation (32) implies that the extra UMSE of CM-minimizing estimators 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 (31) as a function of $J_{u,\nu}(\boldsymbol{q}_{m,\nu})/\sigma_s^2$ for various values of $\rho_{\min}$ and $\rho_{\max}$. The second-order approximation based on (32) appears very good for all but the largest values of UMSE.

C. Comments on the CM-UMSE Bounds

1) Implicit Incorporation of $\mathcal{Q}_a$: First, recall that the CM-UMSE bounding procedure incorporated $\mathcal{Q}_a$, the set

$$J_{u,\nu}\big|_{c,\nu}^{\max, J_{u,\nu}(\boldsymbol{q}_{m,\nu})} := \begin{cases} \dfrac{1-\sqrt{(1+\rho_{\min})\left(1+\frac{J_{u,\nu}(\boldsymbol{q}_{m,\nu})}{\sigma_s^2}\right)^{-2}-\rho_{\min}}}{\rho_{\min}+\sqrt{(1+\rho_{\min})\left(1+\frac{J_{u,\nu}(\boldsymbol{q}_{m,\nu})}{\sigma_s^2}\right)^{-2}-\rho_{\min}}}\,\sigma_s^2, & \kappa_s^{\max} \le \kappa_g \\ \dfrac{1-\sqrt{(1+\rho_{\min})\left(1+\frac{J_{u,\nu}(\boldsymbol{q}_{m,\nu})}{\sigma_s^2}\right)^{-2}\left(1+\rho_{\max}\frac{J_{u,\nu}^2(\boldsymbol{q}_{m,\nu})}{\sigma_s^4}\right)-\rho_{\min}}}{\rho_{\min}+\sqrt{(1+\rho_{\min})\left(1+\frac{J_{u,\nu}(\boldsymbol{q}_{m,\nu})}{\sigma_s^2}\right)^{-2}\left(1+\rho_{\max}\frac{J_{u,\nu}^2(\boldsymbol{q}_{m,\nu})}{\sigma_s^4}\right)-\rho_{\min}}}\,\sigma_s^2, & \kappa_s^{\max} > \kappa_g. \end{cases} \tag{31}$$

$$\begin{aligned} \mathcal{E}_{u,\nu}\big|_{c,\nu}^{\max, J_{u,\nu}(\boldsymbol{q}_{m,\nu})} &:= J_{u,\nu}\big|_{c,\nu}^{\max, J_{u,\nu}(\boldsymbol{q}_{m,\nu})} - J_{u,\nu}(\boldsymbol{q}_{m,\nu}) \\ &= \begin{cases} \frac{1}{2\sigma_s^2}\rho_{\min} J_{u,\nu}^2(\boldsymbol{q}_{m,\nu}) + \mathcal{O}\left(J_{u,\nu}^3(\boldsymbol{q}_{m,\nu})\right), & \kappa_s^{\max} \le \kappa_g \\ \frac{1}{2\sigma_s^2}(\rho_{\min}-\rho_{\max}) J_{u,\nu}^2(\boldsymbol{q}_{m,\nu}) + \mathcal{O}\left(J_{u,\nu}^3(\boldsymbol{q}_{m,\nu})\right), & \kappa_s^{\max} > \kappa_g. \end{cases} \end{aligned} \tag{32}$$

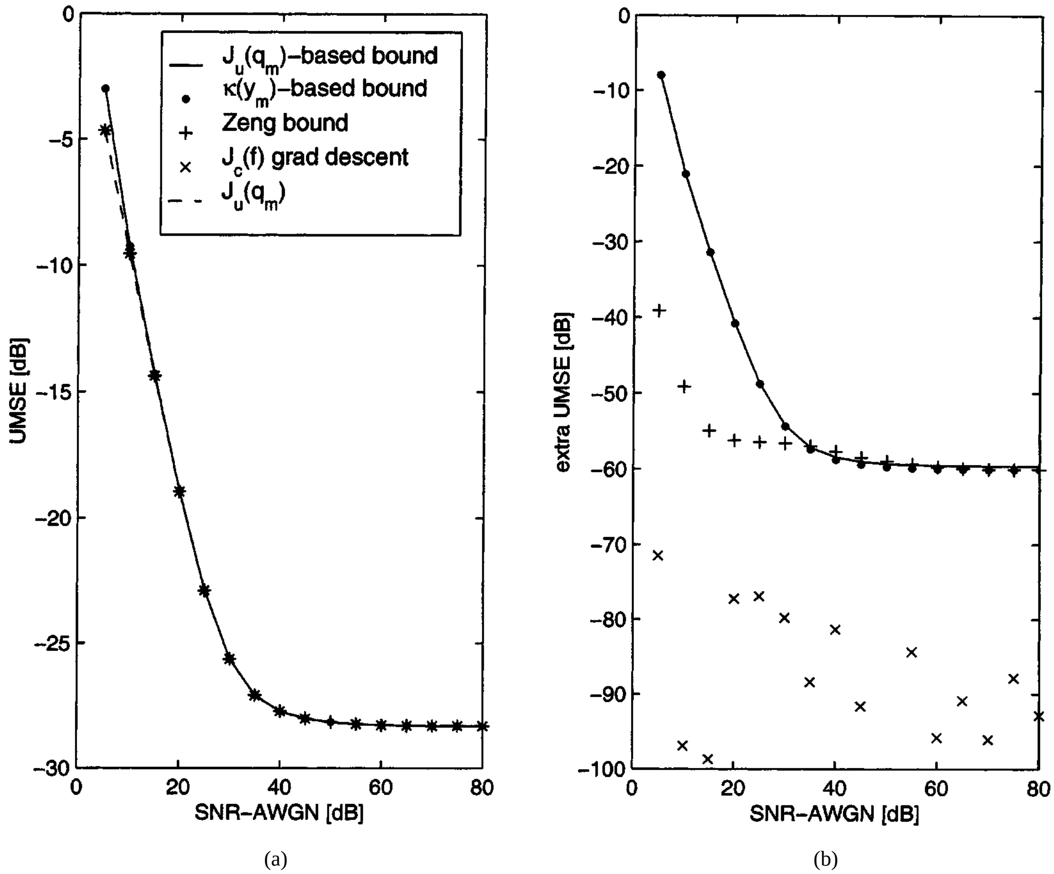


Fig. 5. Bounds on CM-UMSE versus SNR of AWGN for SPIB microwave channel #5, $N_f = 20$, and 8-PAM.

of attainable global responses, *only* in the requirement that $\boldsymbol{q}_{r,\nu} \in \mathcal{Q}_a$. Thus Theorems 1–3, written under the reference choice $\boldsymbol{q}_{r,\nu} = \boldsymbol{q}_{m,\nu} \in \mathcal{Q}_a \cap \mathcal{Q}_\nu^{(0)}$, implicitly incorporate the channel and/or estimator constraints that define $\mathcal{Q}_a$. For example, if $\boldsymbol{q}_{m,\nu}$ is the MMSE response constrained to the set of ARMA estimators, then CM-UMSE bounds based on this $\boldsymbol{q}_{m,\nu}$ will implicitly incorporate the causality constraint. The implicit incorporation of the attainable set $\mathcal{Q}_a$ makes these bounding theorems quite general and easy to use.

2) Effect of $\rho_{\min}$: When

$$\kappa_s^{\max} \le \kappa_g$$

and

$$\rho_{\min} = (\kappa_g - \kappa_s^{\min})/(\kappa_g - \kappa_s^{(0)}) = 1$$

the expressions in Theorems 1–3 simplify

$$\begin{aligned} J_{u,\nu}(\boldsymbol{q}_{c,\nu}) &\le \frac{1 - \sqrt{2\frac{\kappa_g - \kappa_{y_m}}{\kappa_g - \kappa_s^{(0)}} - 1}}{1 + \sqrt{2\frac{\kappa_g - \kappa_{y_m}}{\kappa_g - \kappa_s^{(0)}} - 1}} \sigma_s^2, \\ &\quad \text{when} \quad \frac{1}{2} < \frac{\kappa_g - \kappa_{y_m}}{\kappa_g - \kappa_s^{(0)}} \le 1 \\ &\le \frac{1 - \sqrt{2\left(1 + \frac{J_{u,\nu}(\boldsymbol{q}_{m,\nu})}{\sigma_s^2}\right)^{-2} - 1}}{1 + \sqrt{2\left(1 + \frac{J_{u,\nu}(\boldsymbol{q}_{m,\nu})}{\sigma_s^2}\right)^{-2} - 1}} \sigma_s^2, \\ &\quad \text{when} \quad \frac{J_{u,\nu}(\boldsymbol{q}_{m,\nu})}{\sigma_s^2} < \sqrt{2} - 1 \\ &= J_{u,\nu}(\boldsymbol{q}_{m,\nu}) + \frac{1}{2\sigma_s^2} J_{u,\nu}^2(\boldsymbol{q}_{m,\nu}) \\ &\quad + \mathcal{O}\left(J_{u,\nu}^3(\boldsymbol{q}_{m,\nu})\right). \end{aligned}$$

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

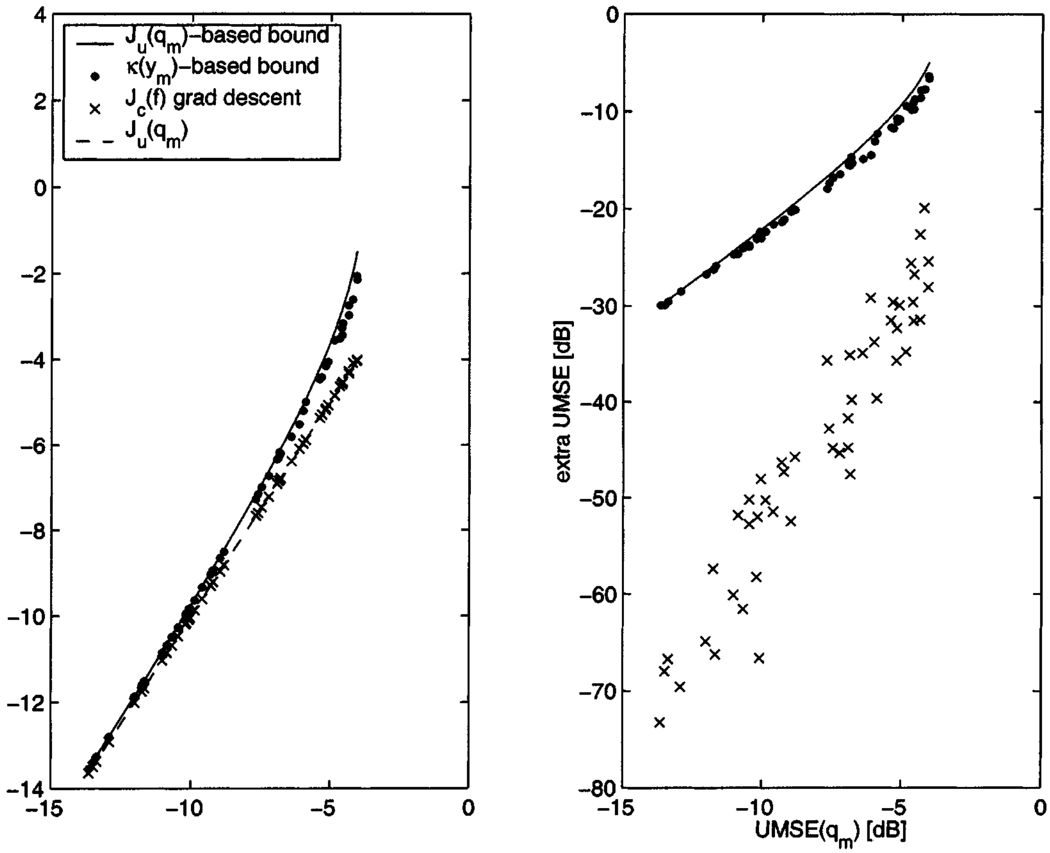
a) sub-Gaussian desired source in the presence of AWGN, or
b) constant-modulus desired source in the presence of non-super-Gaussian interference.

Note that in the two cases above, the CM-UMSE upper bound is independent of the specific distribution of the desired and interfering sources, respectively.

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 $\boldsymbol{q}_{m,\nu}$ via J_o become more stringent (recall Fig. 3).

3) Generalization of Perfect CM-Estimation Property: Finally, we note that the $J_{u,\nu}(\boldsymbol{q}_{m,\nu})$-based CM-UMSE bound in Theorem 2 implies that the perfect CM-estimation property, proven under more restrictive conditions in [16]–[18], extends to the general multisource linear model of Fig. 1.

Corollary 1: CM-minimizing estimators are perfect (up to a scaling) when Wiener estimators are perfect.



(a) (b)

Fig. 6. Bounds on CM-UMSE for $N_f = 8, 10$ BPSK sources, AWGN at -20 dB, and random $\mathcal{H}$.

Proof: From Theorem 2

$$J_{u,\nu}(\boldsymbol{q}_{m,\nu}) = 0 \Rightarrow J_{u,\nu}(\boldsymbol{q}_{c,\nu}) = 0.$$

Hence, the estimators are perfect up to a (fixed) scale factor. □

IV. Numerical Examples

In Sections IV-A–IV.C, we compare the UMSE bounds in (29) and (31) to the UMSE bound of the Zeng *et al.*method of Table I, to the UMSE of the CM estimators found by gradient descent,[4] and to the minimum UMSE (i.e., that obtained by the MMSE solution). The results suggest that, over a wide range of conditions, i) the CM-UMSE bounds are close to the CM-UMSE found by gradient descent, and ii) the CM-UMSE performance is close to the optimal UMSE performance. In other words, the CM-UMSE bounds are tight, and the CM estimator is robust in a MSE sense.

A. Performance Versus Estimator Length for Fixed Channel

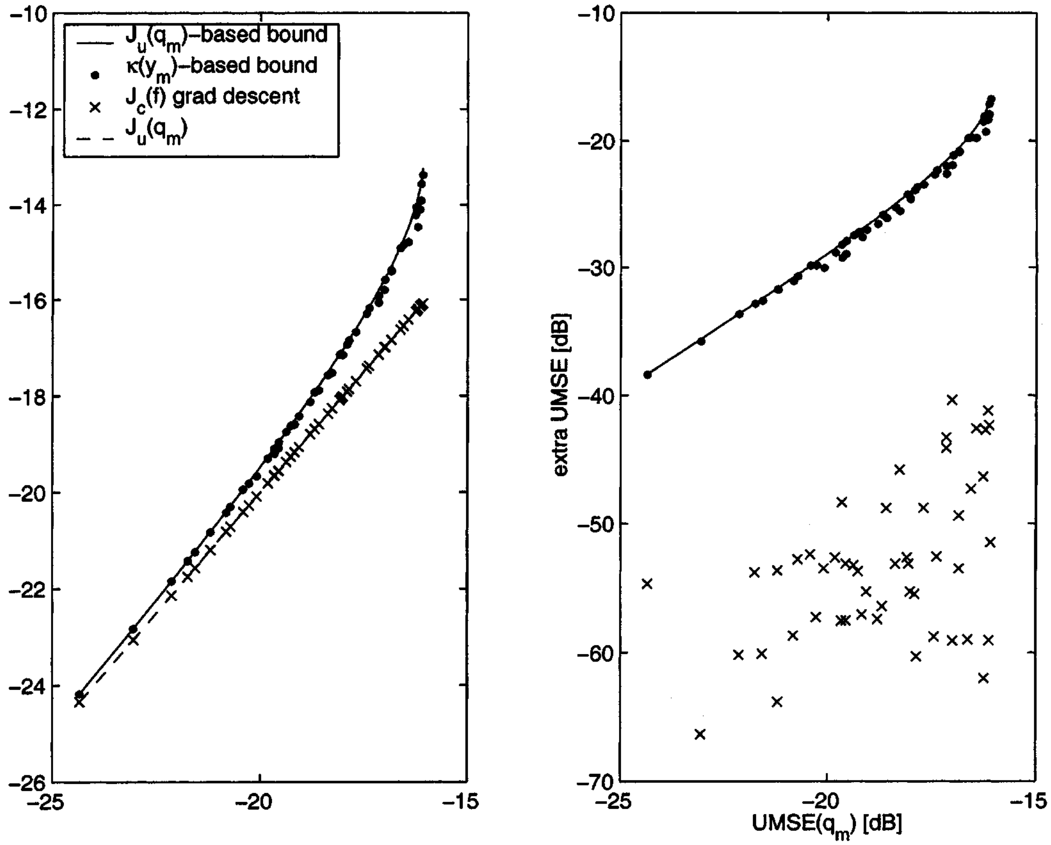
In practical equalization applications, CM-minimizing estimators will not be perfect due to violation of the FCR $\mathcal{H}$ requirement discussed in Section II-D. For instance, even in the absence of noise and interferers, insufficient estimator length can lead to a matrix $\mathcal{H}$ that is wider than tall, thus preventing FCR. For FIR channels with adequate "diversity," it is well known that there exists a finite estimator length sufficient for the achievement of FCR $\mathcal{H}$. When diversity is not adequate, however, as with a baud-spaced scalar channel (i.e., $P = 1$) or with multiple channels sharing common zeros,[5] there exists no finite sufficient length. Consequently, the performance of the CM criterion under so-called "channel undermodeling" and "lack of disparity" has been a topic of recent interest (see, e.g., [8], [11]–[13]).

Using the $T/2$-spaced microwave channel impulse response model #5 from the Signal Processing Information Base (SPIB) database,[6] CM estimator performance was calculated versus estimator length. Fig. 4(a) plots the UMSE of CM-minimizing estimators as predicted by various bounds and by gradient descent. Note that all methods yield CM-UMSE bounds nearly indistinguishable from the minimum UMSE. Fig. 4(b) plots the same information in the form of extra CM-UMSE (i.e., CM-UMSE minus minimum UMSE), and once again we see that the bounds are tight and give nearly identical performance. For the higher equalizer lengths, it is apparent that numerical inaccuracies prevented the CM gradient descent procedure from finding the true minimum (resulting in $\times$'s above the upper bound line).

[4]Gradient descent results were obtained via the MATLAB routine "`fminu`," which was initialized randomly in a small ball around the MMSE estimator.

[5]See, e.g., [6] or [19] for more information on length and diversity requirements.

[6]The SPIB microwave channel database resides at http://spib.rice.edu/spib/microwave.html



(a) (b)

Fig. 7. Bounds on CM-UMSE for $N_f = 8, 5$ BPSK sources, #5 sources with $\kappa_s^{(k)} = 2.9$ (one of which is desired), AWGN at -30 dB, and random $\mathcal{H}$.

B. Performance Versus AWGN for Fixed Channel

Using the same microwave channel model, we conducted a different experiment in which AWGN was introduced at various power levels (for fixed equalizer length $N_f = 20$). Fig. 5(a) shows that the UMSE predicted by the CM bounds is very close to that predicted by gradient descent for all but the highest levels of AWGN, and as before, the CM-UMSE performance is quite close to Weiner UMSE performance. Fig. 5(b) reveals slight differences in bound performance: Zeng *et al.*'s algorithmic bound appears slightly tighter than our closed-form bounds at lower signal-to-noise ratio (SNR).

C. Performance with Random Mixing Matrices

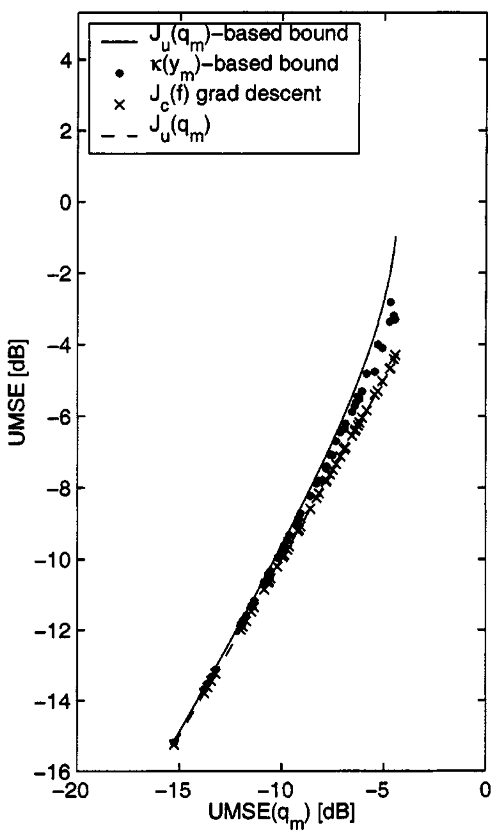
While the convolutive nature of the channel in single-user equalization applications gives $\mathcal{H}$ a block-Toeplitz structure, multiuser applications (e.g., beamforming) may lead to $\mathcal{H}$ with a more general, non-Toeplitz, structure. When the number of sources is greater than the estimator length (which, in our model, is always the case when noise is present), the channel matrix $\mathcal{H}$ will be non-FCR and different estimation techniques will yield different levels of performance.

Here, we present the results of experiments where $\mathcal{H}$ was generated with zero-mean Gaussian entries. Fig. 6 corresponds to a desired source having constant modulus (i.e., $\kappa_s^{(0)} = 1$) in the presence of AWGN and constant modulus interference, Fig. 7 corresponds to a nearly Gaussian desired source in the same interference environment, and Fig. 8 corresponds to a desired source with constant modulus in the presence of AWGN and super-Gaussian interference. As with our previous experiments, Figs. 6–8 demonstrate that i) the closed-form CM-UMSE bounds are tight and ii) that the CM estimators generate nearly MMSE estimates under arbitrary forms of additive interference.

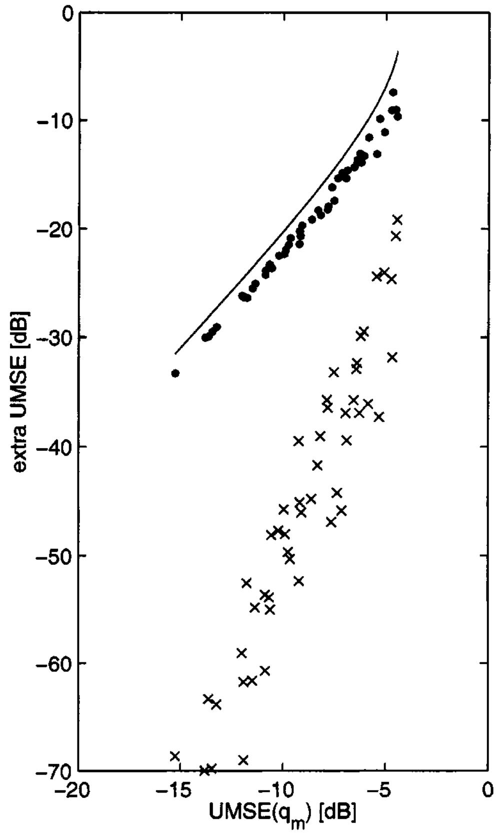
V. Conclusion

In this paper we have derived, for the general multisource linear model of Fig. 1, two closed-form bounding expressions for the UMSE of CM-minimizing estimators. The first bound is based on the kurtosis of the MMSE estimates, while the second is based on the UMSE of the MMSE estimators. Analysis of the second bound shows that the *extra* UMSE of CM-minimizing estimators is upper-bounded by approximately the *square* of the minimum UMSE. Thus the CM-minimizing estimator generates nearly MMSE estimates when the minimum MSE is small. Numerical simulations suggest that the bounds are tight (with respect to the performance of CM-minimizing estimators designed by gradient descent).

This work confirms the long-standing conjecture (see, e.g., [4] and [5]) that the MSE performance of the CM-minimizing estimator is robust to *general* linear channels and general (multisource) additive interference. As such, our results supersede previous work demonstrating the MSE robustness of CM-minimizing estimators in special cases (e.g., when only AWGN is present, when the channel does not provide adequate diversity,



(a)



(b)

Fig. 8. Bounds on CM-UMSE for $N_f = 8$, five BPSK sources (one of which is desired), #5 sources with $\kappa_s^{(k)} = 4$, AWGN at -20 dB, and random $\boldsymbol{\mathcal{H}}$.

or when the estimator has an insufficient number of adjustable parameters).

APPENDIX I
DERIVATION DETAILS FOR CM-UMSE BOUNDS

This appendix contains the proofs of the theorems and lemmas found in Section III.

A. Proof of Lemma 1

In this section we derive an expression for the CM cost J_c in terms of the global response $\boldsymbol{q}$. From (9) and (2)

$$\begin{aligned} J_c(y_n) &= E\{|y_n|^4\} - 2\gamma E\{|y_n|^2\} + \gamma^2 \\ &= \mathcal{K}(y_n) + 2E^2\{|y_n|^2\} + \left|E\left\{y_n^2\right\}\right|^2 - 2\gamma E\left\{|y_n|^2\right\} + \gamma^2. \end{aligned} \tag{33}$$

Source assumptions S1)–S2) imply [20]

$$\mathcal{K}(y_n) = \sum_k \left\|\boldsymbol{q}^{(k)}\right\|_4^4 \mathcal{K}\left(s_n^{(k)}\right). \tag{34}$$

From S3), S5), and the definitions of $\kappa_s^{(k)}$ and κ_g in (19) and (20)

$$\begin{aligned} \mathcal{K}(s_n^{(k)}) &= \begin{cases} E\left\{\left|s_n^{(k)}\right|^4\right\} - 3\sigma_s^4, & \text{real-valued } \{s_n^{(k)}\} \\ E\left\{\left|s_n^{(k)}\right|^4\right\} - 2\sigma_s^4, & E\left\{{s_n^{(k)}}^2\right\} = 0 \end{cases} \\ &= E\left\{\left|s_n^{(k)}\right|^4\right\} - \kappa_g \sigma_s^4 \\ &= \left(\kappa_s^{(k)} - \kappa_g\right)\sigma_s^4. \end{aligned} \tag{35}$$

Similarly, S1)–S3) and S5) imply

$$E\{|y_n|^2\} = \sum_k \|\boldsymbol{q}^{(k)}\|_2^2 \sigma_s^2 = \|\boldsymbol{q}\|_2^2 \sigma_s^2 \tag{36}$$

$$E\left\{y_n^2\right\} = \begin{cases} \|\boldsymbol{q}\|_2^2 \sigma_s^2, & \text{real-valued } \left\{s_n^{(k)}\right\} \\ 0, & E\left\{{s_n^{(k)}}^2\right\} = 0 \quad \forall k. \end{cases} \tag{37}$$

Plugging (34)–(37) into (33), we arrive at (25).

B. Proof of Lemma 2

In this section, we are interested in computing

$$\beta_r = \arg\min_\beta J_c(\beta \boldsymbol{q}_{r,\nu}).$$

For any $\boldsymbol{q}_{r,\nu}$, (25) implies

$$\begin{aligned} \frac{J_c(\beta \boldsymbol{q}_{r,\nu})}{\sigma_s^4} &= \beta^4 \sum_k \left(\kappa_s^{(k)} - \kappa_g\right) \left\|\boldsymbol{q}_{r,\nu}^{(k)}\right\|_4^4 \\ &\quad + \beta^4 \kappa_g \|\boldsymbol{q}_{r,\nu}\|_2^4 - 2\beta^2 \left(\gamma/\sigma_s^2\right) \|\boldsymbol{q}_{r,\nu}\|_2^2 + \left(\gamma/\sigma_s^2\right)^2. \end{aligned} \tag{38}$$

Calculating the first and second partial derivatives of (38) w.r.t. β, it can be shown that

$$\beta_r = \frac{1}{\|\boldsymbol{q}_{r,\nu}\|_2} \sqrt{\left(\frac{\gamma}{\sigma_s^2}\right) \kappa_{y_r}^{-1}}$$

is the minimizing value of β, where

$$\kappa_{y_r} := \left.\frac{E\{|y_n|^4\}}{\sigma_y^4}\right|_{\boldsymbol{q}_{r,\nu}} = \sum_k \left(\kappa_s^{(k)} - \kappa_g\right) \frac{\left\|\boldsymbol{q}_{r,\nu}^{(k)}\right\|_4^4}{\|\boldsymbol{q}_{r,\nu}\|_2^4} + \kappa_g \tag{39}$$

is the normalized kurtosis of the reference estimates. Plugging the expression for β_r into (38), we arrive at the expression for $J_c(\beta_r \boldsymbol{q}_{r,\nu})$ given in (26).

C. Proof of Leema 3

In this section, we are interested in deriving an expression for the interference radius $b_*(a)$ defined in (17) and establishing conditions under which this radius is well defined. Rather than working with (17) directly, we find it easier to use the equivalent definition

$$b_*(a) = \min b(a) \text{ s.t. } \{\|\bar{\boldsymbol{q}}\|_2 > b(a) \Rightarrow J_c(a, \bar{\boldsymbol{q}}) > J_c(\beta_r \boldsymbol{q}_{r,\nu})\}. \tag{40}$$

First we rewrite the CM cost expression (25) in terms of gain $a = \|\boldsymbol{q}\|_2$ and interference response $\bar{\boldsymbol{q}}$ (defined in Section II-C). Using the fact that $|q_\nu^{(0)}|^2 = a^2 - \|\bar{\boldsymbol{q}}\|_2^2$

$$\begin{aligned}
&\sum_k \left(\kappa_s^{(k)} - \kappa_g\right) \left\|\boldsymbol{q}^{(k)}\right\|_4^4 \\
&\quad= \left(\kappa_s^{(0)} - \kappa_g\right) \left|q_\nu^{(0)}\right|^4 \\
&\qquad+ \sum_k \left(\kappa_s^{(k)} - \kappa_g\right) \left\|\bar{\boldsymbol{q}}^{(k)}\right\|_4^4 \\
&\quad= \left(\kappa_s^{(0)} - \kappa_g\right) \left(a^4 - 2a^2\|\bar{\boldsymbol{q}}\|_2^2 + \|\bar{\boldsymbol{q}}\|_2^4\right) \\
&\qquad+ \sum_k \left(\kappa_s^{(k)} - \kappa_g\right) \left\|\bar{\boldsymbol{q}}^{(k)}\right\|_4^4.
\end{aligned}$$

Plugging the previous expression into (25), we find that

$$\begin{aligned}
\frac{J_c(a, \bar{\boldsymbol{q}})}{\sigma_s^4} &= \sum_k \left(\kappa_s^{(k)} - \kappa_g\right) \left\|\bar{\boldsymbol{q}}^{(k)}\right\|_4^4 + \kappa_s^{(0)} a^4 \\
&\quad - 2\left(\kappa_s^{(0)} - \kappa_g\right) a^2 \|\bar{\boldsymbol{q}}\|_2^2 + \left(\kappa_s^{(0)} - \kappa_g\right) \|\bar{\boldsymbol{q}}\|_2^4 \\
&\quad - 2\left(\gamma/\sigma_s^2\right) a^2 + \left(\gamma/\sigma_s^2\right)^2.
\end{aligned} \tag{41}$$

From (26) and (41), the following statements (see (42) at the bottom of this page) are equivalent. The reversal of inequality in (42) occurs because $\kappa_s^{(0)} - \kappa_g < 0$ (as implied by S4)). Since (21) defined $\kappa_s^{\min} = \min_{0 \le k \le K} \kappa_s^{(k)}$, we know that $\kappa_s^{\min} - \kappa_g < 0$. Combining this with the fact that $0 \le \|\bar{\boldsymbol{q}}^{(k)}\|_4^4 / \|\bar{\boldsymbol{q}}^{(k)}\|_2^4 \le 1$, we have

$$\begin{aligned}
\sum_k \left(\kappa_s^{(k)} - \kappa_g\right) \left\|\bar{\boldsymbol{q}}^{(k)}\right\|_4^4 &\ge \left(\kappa_s^{\min} - \kappa_g\right) \sum_k \left\|\bar{\boldsymbol{q}}^{(k)}\right\|_4^4 \\
&\ge \left(\kappa_s^{\min} - \kappa_g\right) \sum_k \left\|\bar{\boldsymbol{q}}^{(k)}\right\|_2^4 \\
&= \left(\kappa_s^{\min} - \kappa_g\right) \|\bar{\boldsymbol{q}}\|_2^4.
\end{aligned}$$

Thus the following is a sufficient condition for (42):

$$0 > \left(1 + \underbrace{\frac{\kappa_s^{\min} - \kappa_g}{\kappa_s^{(0)} - \kappa_g}}_{\rho_{\min}}\right) \|\bar{\boldsymbol{q}}\|_2^4 - 2a^2 \|\bar{\boldsymbol{q}}\|_2^2 + C(a, \boldsymbol{q}_{r,\nu}). \tag{43}$$

Because $1 + \rho_{\min}$ is positive, the set of $\{\|\bar{\boldsymbol{q}}\|_2^2\}$ that satisfy (43) is equivalent to the set of points $\{x\}$ that lie between the roots $\{x_1, x_2\}$ of the quadratic

$$P_a(x) = (1 + \rho_{\min}) x^2 - 2a^2 x + C(a, \boldsymbol{q}_{r,\nu}).$$

Because $\bar{\boldsymbol{q}}$ is an interference response, not all values of $\|\bar{\boldsymbol{q}}\|_2$ are valid. As explained below, we only need to concern ourselves with $0 \le \|\bar{\boldsymbol{q}}\|_2 < a\sqrt{2}^{-1}$. This implies that a valid upper bound on ${b_*}^2(a)$ from (17) is given by the smaller root of $P_a(x)$ when i) this smaller root is nonnegative real and ii) the larger root of $P_a(x)$ is $\ge a^2/2$.

When both roots of $P_a(x)$ lie in the interval $[0, a^2/2)$, there exist two valid regions in the gain-a interference space with CM cost smaller than at the reference, i.e., the set $\{\bar{\boldsymbol{q}} : (a, \bar{\boldsymbol{q}}) \in \mathcal{Q}_c(\beta_r \boldsymbol{q}_{r,\nu})\}$ becomes disjoint. The "inner" part of this disjoint set allows UMSE bounding since it can be contained by $\{\bar{\boldsymbol{q}} : \|\bar{\boldsymbol{q}}\|_2 \le b_1(a)\}$ for a positive interference radius $b_1(a)$, but the "outer" part of the set does *not* permit practial bounding. Such disjointness of $\mathcal{Q}_c(\beta_r \boldsymbol{q}_{r,\nu})$ arises from a source $k \ne 0$ such that $\kappa_s^{(k)} < \kappa_s^{(0)}$. In these scenarios, the point of lowest CM cost in the "outer" regions of $\{\bar{\boldsymbol{q}} : (a, \bar{\boldsymbol{q}}) \in \mathcal{Q}_c(\beta_r \boldsymbol{q}_{r,\nu})\}$ occurs at points on the boundary of

$$\begin{aligned}
&J_c(\beta_r \boldsymbol{q}_{r,\nu}) < J_c(a, \bar{\boldsymbol{q}}) \\
&\quad 0 < \sum_k \left(\kappa_s^{(k)} - \kappa_g\right) \left\|\bar{\boldsymbol{q}}^{(k)}\right\|_4^4 + \left(\kappa_s^{(0)} - \kappa_g\right) \left(-2a^2 \|\bar{\boldsymbol{q}}\|_2^2 + \|\bar{\boldsymbol{q}}\|_2^4\right) \\
&\qquad + \kappa_s^{(0)} a^4 - 2\left(\frac{\gamma}{\sigma_s^2}\right) a^2 + \left(\frac{\gamma}{\sigma_s^2}\right)^2 \kappa_{y_r}^{-1} \\
&\quad 0 > \frac{1}{\kappa_s^{(0)} - \kappa_g} \sum_k \left(\kappa_s^{(k)} - \kappa_g\right) \left\|\bar{\boldsymbol{q}}^{(k)}\right\|_4^4 - 2a^2 \|\bar{\boldsymbol{q}}\|_2^2 + \|\bar{\boldsymbol{q}}\|_2^4 \\
&\qquad + \underbrace{\frac{1}{\kappa_s^{(0)} - \kappa_g} \left(\kappa_s^{(0)} a^4 - 2\left(\frac{\gamma}{\sigma_s^2}\right) a^2 + \left(\frac{\gamma}{\sigma_s^2}\right)^2 \kappa_{y_r}^{-1}\right)}_{C(a, \boldsymbol{q}_{r,\nu})}
\end{aligned} \tag{42}$$

$\mathcal{Q}_\nu^{(0)}$ of the form $\bar{\boldsymbol{q}} = (\ldots, 0, 0, ae^{j\theta}\sqrt{2}^{-1}, 0, 0, \ldots)^t$ and hence with $\|\bar{\boldsymbol{q}}\|_2^2 = a^2/2$. Thus when $x_2 \geq a^2/2$, we can be assured that all valid interference responses (i.e., $\{\bar{\boldsymbol{q}} : (a, \bar{\boldsymbol{q}}) \in \mathcal{Q}_\nu^{(0)}\}$) with CM cost less than the reference can be bounded by some radius b_1.

Solving for the roots of $P_a(x)$ yields (for $x_1 \leq x_2$)

$$\{x_1, x_2\} = \frac{a^2 \pm \sqrt{a^4 - (\rho_{\min}+1)C(a, \boldsymbol{q}_{r,\nu})}}{\rho_{\min}+1} = a^2 \left(\frac{1 \pm \sqrt{1 - (\rho_{\min}+1)\frac{C(a,\boldsymbol{q}_{r,\nu})}{a^4}}}{\rho_{\min}+1} \right)$$

and both roots are nonnegative real when

$$0 \leq C(a, \boldsymbol{q}_{r,\nu})/a^4 \leq (\rho_{\min}+1)^{-1}.$$

It can be shown that $x_2 > a^2/2$ occurs when

$$C(a, \boldsymbol{q}_{r,\nu})/a^4 \leq (3 - \rho_{\min})/4.$$

Since $\rho_{\min} \geq 1$ implies

$$(3 - \rho_{\min})/4 \leq (\rho_{\min}+1)^{-1}$$

both root requirements are satisfied when

$$0 \leq \frac{C(a, \boldsymbol{q}_{r,\nu})}{a^4} \leq \frac{3 - \rho_{\min}}{4}. \tag{44}$$

D. Proof of Theorem 1

In this section we use the expression for $b_*(a)$ from (27) and a suitably chosen reference response $\boldsymbol{q}_{r,\nu} \in \mathcal{Q}_a \cap \mathcal{Q}_\nu^{(0)}$ to upper-bound $J_{u,\nu}(\boldsymbol{q}_{c,\nu})$. Plugging (27) in (18)

$$\frac{J_{u,\nu}(\boldsymbol{q}_{c,\nu})}{\sigma_s^2} \leq \max_a \frac{1 - \sqrt{1 - (\rho_{\min}+1)\frac{C(a,\boldsymbol{q}_{r,\nu})}{a^4}}}{\rho_{\min} + \sqrt{1 - (\rho_{\min}+1)\frac{C(a,\boldsymbol{q}_{r,\nu})}{a^4}}}, \quad \text{when } 0 \leq \frac{C(a, \boldsymbol{q}_{r,\nu})}{a^4} \leq \frac{3 - \rho_{\min}}{4}. \tag{45}$$

Note that the fraction on the right of (45) is nonnegative and strictly increasing in $C(a, \boldsymbol{q}_{r,\nu})/a^4$ over the valid range of $C(a, \boldsymbol{q}_{r,\nu})/a^4$. Hence, finding a that maximizes this expression can be accomplished by finding a that maximizes $C(a, \boldsymbol{q}_{r,\nu})/a^4$. To find these maxima, we first rewrite $C(a, \boldsymbol{q}_{r,\nu})/a^4$ from (42)

$$\frac{C(a, \boldsymbol{q}_{r,\nu})}{a^4} = C_1 \left(\frac{1}{2}(a^2)^{-2} - \kappa_{y_r} \left(\frac{\gamma}{\sigma_s^2} \right)^{-1} (a^2)^{-1} + C_2 \right)$$

where C_1 and C_2 are independent of a. Computing the partial derivative with respect to the quantity a^2

$$\frac{\partial}{\partial(a^2)} \left\{ \frac{C(a, \boldsymbol{q}_{r,\nu})}{a^4} \right\} = C_1 (a^2)^{-3} \left(\kappa_{y_r} \left(\frac{\gamma}{\sigma_s^2} \right)^{-1} a^2 - 1 \right).$$

Setting the partial derivative to zero yields the unique finite maximum

$$a_{\max}^2 = \left(\frac{\gamma}{\sigma_s^2} \right) \kappa_{y_r}.$$

Plugging $a_{\max}^2$ into (42) gives the simple result

$$\frac{C(a_{\max}, \boldsymbol{q}_{r,\nu})}{a_{\max}^4} = \frac{\kappa_s^{(0)} - \kappa_{y_r}}{\kappa_s^{(0)} - \kappa_g} = 1 - \frac{\kappa_g - \kappa_{y_r}}{\kappa_g - \kappa_s^{(0)}}$$

and the $C(a, \boldsymbol{q}_{r,\nu})/a^4$ requirement (44) translates into

$$\frac{1 + \rho_{\min}}{4} \leq \frac{\kappa_g - \kappa_{y_r}}{\kappa_g - \kappa_s^{(0)}} \leq 1. \tag{46}$$

Finally, plugging $C(a_{\max}, \boldsymbol{q}_{r,\nu})/a_{\max}^4$ into (45) gives

$$\frac{J_{u,\nu}(\boldsymbol{q}_{c,\nu})}{\sigma_s^2} \leq \frac{1 - \sqrt{(1+\rho_{\min})\frac{\kappa_g - \kappa_{y_r}}{\kappa_g - \kappa_s^{(0)}} - \rho_{\min}}}{\rho_{\min} + \sqrt{(1+\rho_{\min})\frac{\kappa_g - \kappa_{y_r}}{\kappa_g - \kappa_s^{(0)}} - \rho_{\min}}}. \tag{47}$$

We now establish the existence of an attainable CM-minimizing global response associated with the desired user at delay ν, i.e., $\boldsymbol{q}_{c,\nu} \in \mathcal{Q}_a \cap \mathcal{Q}_\nu^{(0)}$. For simplicity, we assume that the space $\boldsymbol{q}$ is finite-dimensional. We will exploit the Weierstrass theorem ([21, p. 40]), which says that a continuous cost functional has a local minimum in a compact set if there exist points in the interior of the set which give cost lower than anywhere on the boundary.

By definition, all points in $\mathcal{Q}_c(\beta_r \boldsymbol{q}_{r,\nu})$ have CM cost less than or equal to $J_c(y_r)$, the CM cost everywhere on the boundary of $\mathcal{Q}_c(\beta_r \boldsymbol{q}_{r,\nu})$. To make this inequality strict, we expand $\mathcal{Q}_c(\beta_r \boldsymbol{q}_{r,\nu})$ to form the new set $\mathcal{Q}_c'(\beta_r \boldsymbol{q}_{r,\nu})$, defined in terms of boundary cost $J_c(y_r) + \epsilon$ (for arbitrarily small $\epsilon > 0$). Thus all points on the boundary of $\mathcal{Q}_c'(\beta_r \boldsymbol{q}_{r,\nu})$ will have CM cost strictly greater than $J_c(y_r)$. But how do we know that such a set $\mathcal{Q}_c'(\beta_r \boldsymbol{q}_{r,\nu})$ exists? We simply need to reformulate (42) with ϵ-larger $J_c(y_r)$, resulting in ϵ-larger $C(a, \boldsymbol{q}_{r,\nu})$ and a modified quadratic $P_a(x)$ in sufficient condition (43). As long as the new roots (call them x_1' and x_2') satisfy $x_1' \in [0, a^2/2)$ and $x_2' > a^2/2$, the set $\{\bar{\boldsymbol{q}} : (a, \bar{\boldsymbol{q}}) \in \mathcal{Q}_c'(\beta_r \boldsymbol{q}_{r,\nu})\}$ is well-defined, and as long as this holds for the worst case a (i.e., $a_{\max}$), $\mathcal{Q}_c'(\beta_r \boldsymbol{q}_{r,\nu})$ will itself be well-defined. This behavior can be guaranteed, for arbitrarily small ϵ, by replacing (46) with the stricter condition

$$\frac{1 + \rho_{\min}}{4} < \frac{\kappa_g - \kappa_{y_r}}{\kappa_g - \kappa_s^{(0)}} \leq 1. \tag{48}$$

To summarize, (48) guarantees the existence of a closed and bounded set $\mathcal{Q}_c'(\beta_r \boldsymbol{q}_{r,\nu})$ containing an interior point $\beta_r \boldsymbol{q}_{r,\nu}$ with CM cost strictly smaller than all points on the set boundary.

Due to attainibility requirements, our local minimum search must be constrained to the relative interior of the $\mathcal{Q}_a$ manifold (which has been embedded in a possibly higher dimensional $\boldsymbol{q}$-space). Can we apply the Weierstrass theorem on this manifold? First, we know the $\mathcal{Q}_a$ manifold intersects $\mathcal{Q}_c'(\beta_r \boldsymbol{q}_{r,\nu})$, namely, at the point $\beta_r \boldsymbol{q}_{r,\nu}$. Second, we know that the relative boundary of the $\mathcal{Q}_a$ manifold occurs outside $\mathcal{Q}_c'(\beta_r \boldsymbol{q}_{r,\nu})$, namely, at infinity. These two observations imply that the boundary of $\mathcal{Q}_a \cap \mathcal{Q}_c'(\beta_r \boldsymbol{q}_{r,\nu})$ relative to $\mathcal{Q}_a$ must be a subset of the boundary of $\mathcal{Q}_c'(\beta_r \boldsymbol{q}_{r,\nu})$. Hence, the interior of $\mathcal{Q}_a \cap \mathcal{Q}_c'(\beta_r \boldsymbol{q}_{r,\nu})$ relative to $\mathcal{Q}_a$ contains points which give kurtosis strictly higher than those on the boundary of $\mathcal{Q}_a \cap \mathcal{Q}_c'(\beta_r \boldsymbol{q}_{r,\nu})$ relative to $\mathcal{Q}_a$. Finally, the domain (i.e., $\mathcal{Q}_a \cap \mathcal{Q}_c'(\beta_r \boldsymbol{q}_{r,\nu})$ relative to $\mathcal{Q}_a$) is closed and bounded, hence compact. Thus the Weierstrass theorem ensures the existence of a local CM minimum in the interior of $\mathcal{Q}_a \cap \mathcal{Q}_c'(\beta_r \boldsymbol{q}_{r,\nu})$ relative to $\mathcal{Q}_a$ under (48). Recalling that $\mathcal{Q}_c'(\beta_r \boldsymbol{q}_{r,\nu}) \subset \mathcal{Q}_\nu^{(0)}$, we see that there exists an attainable locally CM-minimizing response associated with the desired user at delay ν.

Theorem 1 follows directly from (47) with reference choice $\boldsymbol{q}_{r,\nu} = \boldsymbol{q}_{m,\nu} \in \mathcal{Q}_a \cap \mathcal{Q}_\nu^{(0)}$. Note that we restrict ourselves to $\boldsymbol{q}_{m,\nu} \in \mathcal{Q}_\nu$, which may not always be the case.

E. Proof of Theorem 2

In this section we find an upper bound for $J_{u,\nu}(\boldsymbol{q}_{c,\nu})$ that involves the UMSE of a reference estimator $J_{u,\nu}(\boldsymbol{q}_{r,\nu})$ rather than the kurtosis of reference estimates κ_{y_r}. The choice of the reference to be the MMSE estimator can be considered a special case. The conditions we establish below will guarantee that $\boldsymbol{q}_{m,\nu} \in \mathcal{Q}_\nu^{(0)}$.

We will take advantage of the fact that $J_{u,\nu}|_{c,\nu}^{\max,\kappa_{y_r}}$ in (47) is a strictly decreasing function of $(\kappa_g - \kappa_{y_r})/(\kappa_g - \kappa_s^{(0)})$ over its valid range. From (39)

$$\frac{\kappa_g-\kappa_y}{\kappa_g-\kappa_s^{(0)}} = \frac{\sum_k \left(\kappa_g - \kappa_s^{(k)}\right) \left\|\boldsymbol{q}^{(k)}\right\|_4^4}{\left(\kappa_g - \kappa_s^{(0)}\right) \|\boldsymbol{q}\|_2^4} = \frac{\left(\kappa_g-\kappa_s^{(0)}\right) |q_\nu^{(0)}|^4 + \sum_k \left(\kappa_g - \kappa_s^{(k)}\right) \left\|\bar{\boldsymbol{q}}^{(k)}\right\|_4^4}{\left(\kappa_g-\kappa_s^{(0)}\right) \|\boldsymbol{q}\|_2^4}.$$

Examining the previous equation, $0 \le (\|\bar{\boldsymbol{q}}\|_4^4/\|\bar{\boldsymbol{q}}\|_2^4) \le 1$ implies that

$$\sum_k \left(\kappa_g-\kappa_s^{(k)}\right) \left\|\bar{\boldsymbol{q}}^{(k)}\right\|_4^4 \ge (\kappa_g - \kappa_s^{\max}) \|\bar{\boldsymbol{q}}\|_4^4 \ge \begin{cases} 0, & \kappa_s^{\max} \le \kappa_g \\ (\kappa_g-\kappa_s^{\max}) \|\bar{\boldsymbol{q}}\|_2^4, & \kappa_s^{\max} > \kappa_g \end{cases} \tag{49}$$

and

$$\sum_k \left(\kappa_g-\kappa_s^{(k)}\right) \left\|\bar{\boldsymbol{q}}^{(k)}\right\|_4^4 \le \left(\kappa_g-\kappa_s^{\min}\right) \|\bar{\boldsymbol{q}}\|_4^4 \le \left(\kappa_g-\kappa_s^{\min}\right) \|\bar{\boldsymbol{q}}\|_2^4. \tag{50}$$

Note that in (50) and the super-Gaussian case of (49), equality is reached by global responses of the form $\bar{\boldsymbol{q}} = \alpha \boldsymbol{e}_i^{(k)}$, where k corresponds to the source with minimum and maximum kurtosis, respectively.

Considering first the sub-Gaussian interference case $(\kappa_s^{\max} \le \kappa_g)$, we claim

$$\frac{\kappa_g - \kappa_y}{\kappa_g - \kappa_s^{(0)}} \ge \frac{\left|q_\nu^{(0)}\right|^4}{\|\boldsymbol{q}\|_2^4} = \left(1 + \frac{J_{u,\nu}(\boldsymbol{q})}{\sigma_s^2}\right)^{-2} \tag{51}$$

since the definition of $J_{u,\nu}(\boldsymbol{q})$ in (14) implies

$$\|\boldsymbol{q}\|_2^4 = \left(\left|q_\nu^{(0)}\right|^2 + \|\bar{\boldsymbol{q}}\|_2^2\right)^2 = \left|q_\nu^{(0)}\right|^4 \left(1 + \frac{\|\bar{\boldsymbol{q}}\|_2^2}{\left|q_\nu^{(0)}\right|^2}\right)^2 = \left|q_\nu^{(0)}\right|^4 \left(1 + \frac{J_{u,\nu}(\boldsymbol{q})}{\sigma_s^2}\right)^2.$$

Applying (51) to (47), we obtain

$$\frac{J_{u,\nu}|_{c,\nu}^{\max,\kappa_{y_r}}}{\sigma_s^2} \le \frac{1 - \sqrt{(1+\rho_{\min})\left(1 + \frac{J_{u,\nu}(\boldsymbol{q}_{r,\nu})}{\sigma_s^2}\right)^{-2} - \rho_{\min}}}{\rho_{\min} + \sqrt{(1+\rho_{\min})\left(1 + \frac{J_{u,\nu}(\boldsymbol{q}_{r,\nu})}{\sigma_s^2}\right)^{-2} - \rho_{\min}}}$$

when (48) is satisfied. Inequality (51) implies that

$$\frac{1+\rho_{\min}}{4} < \left(1 + \frac{J_{u,\nu}(\boldsymbol{q}_{r,\nu})}{\sigma_s^2}\right)^{-2} \Leftrightarrow \frac{J_{u,\nu}(\boldsymbol{q}_{r,\nu})}{\sigma_s^2} < -1 + \frac{2}{\sqrt{\rho_{\min}+1}}$$

is sufficient for the left inequality of (48). Turning our attention to the right inequality of (48), we can use (50) to say

$$\frac{\kappa_g - \kappa_y}{\kappa_g - \kappa_s^{(0)}} \le \frac{\left|q_\nu^{(0)}\right|^4}{\|\boldsymbol{q}\|_2^4} + \rho_{\min}\frac{\|\bar{\boldsymbol{q}}\|_2^4}{\|\boldsymbol{q}\|_2^4} = \left(1 + \frac{J_{u,\nu}(\boldsymbol{q})}{\sigma_s^2}\right)^{-2} \left(1 + \rho_{\min}\frac{J_{u,\nu}^2(\boldsymbol{q})}{\sigma_s^4}\right) \tag{52}$$

since (14) implies

$$\frac{\|\bar{\boldsymbol{q}}\|_2^4}{\|\boldsymbol{q}\|_2^4} = \left(\frac{\|\bar{\boldsymbol{q}}\|_2^2 + \left|q_\nu^{(0)}\right|^2}{\|\bar{\boldsymbol{q}}\|_2^2}\right)^{-2} = \left(1 + \frac{\sigma_s^2}{J_{u,\nu}(\boldsymbol{q})}\right)^{-2} = \frac{J_{u,\nu}^2(\boldsymbol{q})}{\sigma_s^4}\left(1 + \frac{J_{u,\nu}(\boldsymbol{q})}{\sigma_s^2}\right)^{-2}.$$

Then, inequality (52) implies that a sufficient condition for the right side of (48) is

$$\left(1 + \frac{J_{u,\nu}(\boldsymbol{q}_{r,\nu})}{\sigma_s^2}\right)^{-2} \left(1 + \rho_{\min}\frac{J_{u,\nu}^2(\boldsymbol{q}_{r,\nu})}{\sigma_s^4}\right) \le 1 \Leftrightarrow \frac{J_{u,\nu}(\boldsymbol{q}_{r,\nu})}{\sigma_s^2} \le \frac{2}{\rho_{\min}-1}.$$

Using $\rho_{\min} \ge 1$, it can be shown that

$$-1 + 2/\sqrt{\rho_{\min}+1} \le 2/(\rho_{\min}-1).$$

Thus satisfaction of our sufficient condition for the left inequality in (48) suffices for both inequalities in (48).

Treatment of the super-Gaussian interference case $(\kappa_s^{\max} > \kappa_g)$ is analogous. With the methods used to obtain (52), (49) implies

$$\frac{\kappa_g-\kappa_y}{\kappa_g-\kappa_s^{(0)}} \ge \frac{\left|q_\nu^{(0)}\right|^4}{\|\boldsymbol{q}\|_2^4} + \underbrace{\frac{\kappa_g-\kappa_s^{\max}}{\kappa_g-\kappa_s^{(0)}}}_{\rho_{\max}} \frac{\|\bar{\boldsymbol{q}}\|_2^4}{\|\boldsymbol{q}\|_2^4} = \left(1+\frac{J_{u,\nu}(\boldsymbol{q})}{\sigma_s^2}\right)^{-2} \left(1+\rho_{\max}\frac{J_{u,\nu}^2(\boldsymbol{q})}{\sigma_s^4}\right). \tag{53}$$

Applying (53) to (47), we obtain the first expression at the bottom of this page as long as (48) is satisfied. Substituting $|q_\nu^{(0)}|^2 = \|\boldsymbol{q}\|_2^2 - \|\bar{\boldsymbol{q}}\|_2^2$ in (53), we find that

$$\frac{\kappa_g - \kappa_y}{\kappa_g - \kappa_s^{(0)}} \geq 1 - 2\frac{\|\bar{\boldsymbol{q}}\|_2^2}{\|\boldsymbol{q}\|_2^2} + (1+\rho_{\max})\frac{\|\bar{\boldsymbol{q}}\|_2^4}{\|\boldsymbol{q}\|_2^4}$$

hence a sufficient condition for the left inequality of (48) becomes

$$(1+\rho_{\min})/4 < 1 - 2\|\bar{\boldsymbol{q}}_{r,\nu}\|_2^2/\|\boldsymbol{q}_{r,\nu}\|_2^2 + (1+\rho_{\max})(\|\bar{\boldsymbol{q}}_{r,\nu}\|_2^2/\|\boldsymbol{q}_{r,\nu}\|_2^2)^2$$

or, equivalently,

$$(1+\rho_{\max})\left(\frac{\|\bar{\boldsymbol{q}}_{r,\nu}\|_2^2}{\|\boldsymbol{q}_{r,\nu}\|_2^2}\right)^2 - 2\frac{\|\bar{\boldsymbol{q}}_{r,\nu}\|_2^2}{\|\boldsymbol{q}_{r,\nu}\|_2^2} + (3-\rho_{\min})/4 < 0.$$

It can be shown that the quadratic inequality above is satisfied by

$$\frac{\|\bar{\boldsymbol{q}}_{r,\nu}\|_2^2}{\|\boldsymbol{q}_{r,\nu}\|_2^2} < \begin{cases} \frac{1-\sqrt{1-(3-\rho_{\min})(1+\rho_{\max})/4}}{1+\rho_{\max}}, & \rho_{\max} \neq -1 \\ (3-\rho_{\min})/8, & \rho_{\max} = -1 \end{cases}$$

and since

$$J_{u,\nu}(\boldsymbol{q}) = (\|\bar{\boldsymbol{q}}\|_2^2/\|\boldsymbol{q}\|_2^2)/(1 - \|\bar{\boldsymbol{q}}\|_2^2/\|\boldsymbol{q}\|_2^2)$$

is strictly increasing in $\|\bar{\boldsymbol{q}}\|_2^2\|\boldsymbol{q}\|_2^2$, the following must be sufficient for the left inequality of (48):

$$\frac{J_{u,\nu}(\boldsymbol{q}_{r,\nu})}{\sigma_s^2} < \begin{cases} \frac{1-\sqrt{1-(3-\rho_{\min})(1+\rho_{\max})/4}}{\rho_{\max}+\sqrt{1-(3-\rho_{\min})(1+\rho_{\max})/4}}, & \rho_{\max} \neq -1 \\ \frac{3-\rho_{\min}}{5+\rho_{\min}}, & \rho_{\max} = -1. \end{cases} \tag{54}$$

As for the right inequality of (48), it can be shown that the quantities in (54) are smaller than $2/(\rho_{\min}-1)$. Thus satisfaction of (54) suffices for both inequalities in (48).

F. Proof of Theorem 3

Here, we reformulate the upper bound (31). To simplify the presentation of the proof, the shorthand notation $J := J_{u,\nu}(\boldsymbol{q}_{m,\nu})/\sigma_s^2$ will be used.

Starting with the non-super-Gaussian case (i.e., $\kappa_s^{\max} \leq \kappa_g$), (31) says

$$\frac{J_{u,\nu}\big|_{c,\nu}^{\max,J_{u,\nu}(\boldsymbol{q}_{m,\nu})}}{\sigma_s^2} = \frac{1-\sqrt{(1+\rho_{\min})(1+J)^{-2}-\rho_{\min}}}{\rho_{\min}+\sqrt{(1+\rho_{\min})(1+J)^{-2}-\rho_{\min}}}$$

from which routine manipulations yield

$$\frac{J_{u,\nu}\big|_{c,\nu}^{\max,J_{u,\nu}(\boldsymbol{q}_{m,\nu})}}{\sigma_s^2} = \frac{1-\sqrt{1-((\rho_{\min}-1)+\rho_{\min}(2J+J^2))\,(2J+J^2)}}{(\rho_{\min}-1)+\rho_{\min}(2J+J^2)}.$$

For $x \in \mathbb{R}$ such that $|x| < 1$, the binomial series [22] may be used to claim

$$\sqrt{1-x} = 1 - \frac{x}{2} - \frac{x^2}{8} - \mathcal{O}(x^3).$$

Applying the previous expression with

$$x = ((\rho_{\min}-1) + \rho_{\min}(2J+J^2))(2J+J^2)$$

we find the second expression at the bottom of this page. Finally, subtraction of J gives the first case in (32).

For the super-Gaussian case (i.e., $\kappa_s^{\max} > \kappa_g$), (31) becomes the first expression at the top of the following page from which routine manipulations yield the second expression at the top of the following page. As before, we use the binomial series expansion for $\sqrt{1-x}$, but now with

$$x = (2\rho_{\max}-2)J + (5\rho_{\min}-1-\rho_{\max}-\rho_{\min}\rho_{\max})J^2 + \mathcal{O}(J^3).$$

After some algebra, we find

$$\frac{J_{u,\nu}\big|_{c,\nu}^{\max,J_{u,\nu}(\boldsymbol{q}_{m,\nu})}}{\sigma_s^2} = J + \frac{1}{2}\frac{(\rho_{\min}-\rho_{\max})(\rho_{\min}-1)J^2 + \mathcal{O}(J^3)}{(\rho_{\min}-1)+2\rho_{\min}J+(\rho_{\min}-\rho_{\max})J^2}.$$

Finally, we apply the series approximation

$$\frac{1}{1-y} = 1 + y + \mathcal{O}(y^2)$$

$$\frac{J_{u,\nu}\big|_{c,\nu}^{\max,\kappa_{y_r}}}{\sigma_s^2} \leq \frac{1-\sqrt{(1+\rho_{\min})\left(1+\frac{J_{u,\nu}(\boldsymbol{q}_{r,\nu})}{\sigma_s^2}\right)^{-2}\left(1+\rho_{\max}\frac{J_{u,\nu}^2(\boldsymbol{q}_{r,\nu})}{\sigma_s^4}\right)-\rho_{\min}}}{\rho_{\min}+\sqrt{(1+\rho_{\min})\left(1+\frac{J_{u,\nu}(\boldsymbol{q}_{r,\nu})}{\sigma_s^2}\right)^{-2}\left(1+\rho_{\max}\frac{J_{u,\nu}^2(\boldsymbol{q}_{r,\nu})}{\sigma_s^4}\right)-\rho_{\min}}}$$

$$\begin{aligned}\frac{J_{u,\nu}\big|_{c,\nu}^{\max,J_{u,\nu}(\boldsymbol{q}_{m,\nu})}}{\sigma_s^2} &= \frac{1}{2}\frac{\left((\rho_{\min}-1)+\rho_{\min}(2J+J^2)\right)(2J+J^2)}{(\rho_{\min}-1)+\rho_{\min}(2J+J^2)} \\ &\quad + \frac{1}{8}\frac{\left((\rho_{\min}-1)+\rho_{\min}(2J+J^2)\right)^2(2J+J^2)^2}{(\rho_{\min}-1)+\rho_{\min}(2J+J^2)} + \mathcal{O}(J^3) \\ &= J + \frac{\rho_{\min}}{2}J^2 + \mathcal{O}(J^3).\end{aligned}$$

$$\frac{J_{u,\nu}|_{c,\nu}^{\max, J_{u,\nu}(\boldsymbol{q}_{m,\nu})}}{\sigma_s^2} = \frac{1-\sqrt{(1+\rho_{\min})(1+J)^{-2}(1+\rho_{\max}J^2)-\rho_{\min}}}{\rho_{\min}+\sqrt{(1+\rho_{\min})(1+J)^{-2}(1+\rho_{\max}J^2)-\rho_{\min}}}$$

$$\frac{J_{u,\nu}|_{c,\nu}^{\max, J_{u,\nu}(\boldsymbol{q}_{m,\nu})}}{\sigma_s^2} = \frac{\rho_{\max}J^2+1-\sqrt{1+(2-2\rho_{\max})J+(1+\rho_{\max}+\rho_{\min}\rho_{\max}-5\rho_{\min})J^2+\mathcal{O}(J^3)}}{(\rho_{\min}-1)+2\rho_{\min}J+(\rho_{\min}-\rho_{\max})J^2}.$$

with

$$y = -(2\rho_{\min}J+(\rho_{\min}-\rho_{\max})J^2)/(\rho_{\min}-1)$$

for $\rho_{\min} \neq 1$. Straightforward algebra yields

$$\frac{J_{u,\nu}|_{c,\nu}^{\max, J_{u,\nu}(\boldsymbol{q}_{m,\nu})}}{\sigma_s^2} = J + \frac{1}{2}(\rho_{\min}-\rho_{\max})J^2 + \mathcal{O}(J^3).$$

Taking the limit $\rho_{\min} \to 1$, it is evident that no problems arise at the point $\rho_{\min} = 1$. Subtraction of J from the last statement gives the second case in (32).

REFERENCES

[1] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Trans. Inform. Theory*, vol. 36, pp. 312–321, Mar. 1990.

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

[3] P. Regalia, "On the equivalence between the Godard and Shalvi-Weinstein schemes of blind equalization," *Signal Processing*, vol. 73, no. 1–2, pp. 185–190, Feb. 1999.

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

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

[6] C. R. Johnson, Jr., P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, and R. A. Casas, "Blind equalization using the constant modulus criterion: A review," *Proc. IEEE (Special Issue on Blind System Identification and Estimation)*, vol. 86, pp. 1927–1950, Oct. 1998.

[7] J. J. Shynk and R. P. Gooch, "The constant modulus array for cochannel signal copy and direction finding," *IEEE Trans. Signal Processing*, vol. 44, pp. 652–660, Mar. 1996.

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

[9] H. H. Zeng, L. Tong, and C. R. Johnson Jr., "Relationships between the constant modulus and Wiener receivers," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1523–1538, July 1998.

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

[11] Y. Li, K. J. R. Liu, and Z. Ding, "Length and cost dependent local minima of unconstrained blind channel equalizers," *IEEE Trans. Signal Processing*, vol. 44, pp. 2726–2735, Nov. 1996.

[12] 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. Signal Processing*, vol. 47, pp. 544–549, Feb. 1999.

[13] P. Regalia and M. Mboup, "Undermodeled equalization: A characterization of stationary points for a family of blind criteria," *IEEE Trans. Signal Processing*, vol. 47, pp. 760–770, Mar. 1999.

[14] H. H. Zeng, L. Tong, and C. R. Johnson, Jr., "An analysis of constant modulus receivers," *IEEE Trans. Signal Processing*, vol. 47, pp. 2990–2999, Nov. 1999.

[15] P. Schniter, "Blind estimation without priors: Performance, convergence, and efficient implementation," Ph.D. dissertation, Cornell Univ., Ithaca, NY, May 2000.

[16] G. J. Foschini, "Equalizing without altering or detecting data (digital radio systems)," *AT&T Tech. J.*, vol. 64, no. 8, pp. 1885–1911, Oct. 1985.

[17] L. Tong, "A fractionally spaced adaptive blind equalizer," in *Proc. Conf. Information Science and Systems*, Princeton, NJ, Mar. 1992, pp. 711–716.

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

[19] C. R. Johnson, Jr. *et al.*, "The core of FSE-CMA behavior theory," in *Unsupervised Adaptive Filtering, Volume 2: Blind Deconvolution*, S. Haykin *et al.*, Ed. New York: Wiley, 2000.

[20] B. Porat, *Digital Processing of Random Signals*. Englewood Cliffs, NJ: Prentice-Hall, 1994.

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

[22] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed. New York: McGraw-Hill, 1976.