Existence and Performance of Shalvi-Weinstein Estimators

Phil Schniter (*schniter@ee.eng.ohio-state.edu*)
Dept. of Electrical Engineering, The Ohio State University, Columbus, OH 43210.

Abstract: The Shalvi-Weinstein (SW) criterion has become popular in the design of blind linear estimators of i.i.d. processes transmitted through unknown linear channels in the presence of unknown additive interference. Here we analyze SW estimators in a general multiple-input multiple-output (MIMO) setting that allows near-arbitrary source/interference distributions and non-invertible channels. The main contributions of this paper are (i) a simple test for the existence of a SW estimator for the desired source and (ii) bounding expressions for the MSE of SW estimators that are a function of the minimum attainable MSE and the kurtoses of the source and interferers.

1 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)

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

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).

In [1], Shalvi and Weinstein showed that, for i.i.d. sources, noiseless invertible channels, and adequately parameterized estimators, perfect blind estimation is possible with knowledge of only the second- and fourth-order moments of the estimates $\{y_n\}$. Based on this observation, they proposed a blind estimation scheme that manipulates $\mathbf{f}(z)$ to maximize the absolute kurtosis of the estimates $\{y_n\}$ subject to a variance constraint:

$$\max\big|\mathcal{K}(y_n)\big| \quad \text{s.t.} \quad \sigma_y^2 = 1, \tag{2}$$

where, for zero-mean $\{y_n\}$, we define variance as $\sigma_y^2 := \mathrm{E}\{|y_n|^2\}$ and kurtosis as

$$\mathcal{K}(y_n) \;:=\; \mathrm{E}\{|y_n|^4\} - 2\,\mathrm{E}^2\{|y_n|^2\} - \big|\mathrm{E}\{y_n^2\}\big|^2. \tag{3}$$

As proven in [1], unconstrained linear estimators locally maximizing the Shalvi-Weinstein (SW) criterion yield perfect blind estimates of a single non-Gaussian i.i.d. source transmitted through a noiseless invertible linear channel. In practical situations, however, we expect inadequately parameterized estimators, non-invertible channels, as well as noise and/or interference of a potentially non-Gaussian nature. Are SW estimators useful in these cases? How do SW estimates compare to optimal (linear) estimators, say, in a mean-squared sense?

In this paper we study the performance of constrained ARMA SW estimators under the assumptions of the model in Section 2.1: desired source with arbitrary non-Gaussian distribution, interference with arbitrary distribution, and vector ARMA channels. The main contributions of this paper are (i) a simple test for the existence of a SW estimator for the desired source, and (ii) bounding expressions for the MSE of SW estimators that are a function of the minimum MSE attainable under the same conditions. These bounds, derived under the multi-source linear model of Fig. 1, provide a formal link between the SW and Wiener estimators in a very general context.

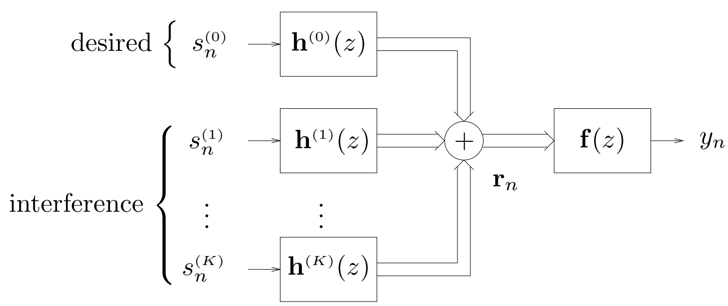

Figure 1: Linear system model with K sources of interference.

2 Background

In this section, we give more detailed information on the linear system model and the MSE, unbiased MSE, and SW criteria. The following notation is used throughout: $(\cdot)^t$ 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 $\sqrt[p]{\sum_i |x_i|^p}$, $\mathbf{I}$ the identity matrix.

2.1 Linear System Model

First we formalize the linear time-invariant multi-channel 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 $\{\mathbf{r}_n\}$ 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. Note that such $\mathbf{h}^{(k)}(z)$ admit infinite impulse response (IIR) channel models.

From the vector-valued observation sequence $\{\mathbf{r}_n\}$, the receiver generates a sequence of linear estimates $\{y_n\}$ of $\{s^{(0)}_{n-\nu}\}$, 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, i.e., certain polynomial coefficients in the numerator and denominator of $\mathbf{f}(z)$ may be held at zero.

In the sequel, we will focus almost exclusively on the global channel-plus-estimator response $q^{(k)}(z) := \mathbf{f}^H(z)\mathbf{h}^{(k)}(z)$. The impulse response coefficients of $q^{(k)}(z)$ can be written $q^{(k)}_n = \sum_{i=-\infty}^{\infty} \mathbf{f}_i^H \mathbf{h}^{(k)}_{n-i}$, allowing the estimates to be written as $y_n = \sum_{k=0}^{K} \sum_{i=-\infty}^{\infty} q^{(k)}_i s^{(k)}_{n-i}$. Adopting the following vector notation helps to streamline the remainder of the paper.

$$\begin{aligned}
\boldsymbol{q}^{(k)} &:= (\ldots, q^{(k)}_{-1}, q^{(k)}_0, q^{(k)}_1, \ldots)^t, \\
\boldsymbol{q} &:= (\cdots,\ q^{(0)}_{-1}, q^{(1)}_{-1}, \ldots, q^{(K)}_{-1},\ q^{(0)}_0, q^{(1)}_0, \ldots, q^{(K)}_0,\ q^{(0)}_1, q^{(1)}_1, \ldots, q^{(K)}_1,\ \cdots)^t, \\
\boldsymbol{s}^{(k)}(n) &:= (\ldots, s^{(k)}_{n+1}, s^{(k)}_n, s^{(k)}_{n-1}, \ldots)^t, \\
\boldsymbol{s}(n) &:= (\cdots,\ s^{(0)}_{n+1}, s^{(1)}_{n+1}, \ldots, s^{(K)}_{n+1},\ s^{(0)}_n, s^{(1)}_n, \ldots, s^{(K)}_n,\ s^{(0)}_{n-1}, s^{(1)}_{n-1}, \ldots, s^{(K)}_{n-1},\ \cdots)^t.
\end{aligned}$$

For instance, the estimates can be rewritten concisely as $y_n = \sum_{k=0}^{K} \boldsymbol{q}^{(k)t}\boldsymbol{s}^{(k)}(n) = \boldsymbol{q}^t\boldsymbol{s}(n)$. The source-specific unit vector $\mathbf{e}^{(k)}_\nu$ will also prove convenient. $\mathbf{e}^{(k)}_\nu$ is a column vector with a single nonzero element of value 1 located such that $\boldsymbol{q}^t\mathbf{e}^{(k)}_\nu = q^{(k)}_\nu$.

We now point out two important properties of $\boldsymbol{q}$. First, it is important to recognize that a particular channel and set of estimator constraints will restrict the set of *attainable* global responses, which we will denote by $\mathcal{Q}_\mathrm{a}$. For example, when the estimator is FIR, $\boldsymbol{q} \in \mathcal{Q}_\mathrm{a} = \mathrm{row}(\boldsymbol{\mathcal{H}})$, where

$$\boldsymbol{\mathcal{H}} := \begin{pmatrix} \mathbf{h}^{(0)}_0 \cdots \mathbf{h}^{(K)}_0 & \mathbf{h}^{(0)}_1 \cdots \mathbf{h}^{(K)}_1 & \mathbf{h}^{(0)}_2 \cdots \mathbf{h}^{(K)}_2 & \cdots \\ \mathbf{0} \ \cdots \ \mathbf{0} & \mathbf{h}^{(0)}_0 \cdots \mathbf{h}^{(K)}_0 & \mathbf{h}^{(0)}_1 \cdots \mathbf{h}^{(K)}_1 & \cdots \\ \vdots \quad\ \ \vdots & \vdots \quad\ \ \vdots & \vdots \quad\ \ \vdots & \\ \mathbf{0} \ \cdots \ \mathbf{0} & \mathbf{0} \ \cdots \ \mathbf{0} & \mathbf{h}^{(0)}_0 \cdots \mathbf{h}^{(K)}_0 & \cdots \end{pmatrix}. \qquad (4)$$

Restricting the estimator to be sparse or autoregressive, for example, would generate different attainable sets $\mathcal{Q}_\mathrm{a}$. Next, BIBO stable $\mathbf{f}(z)$ and $\mathbf{h}^{(k)}(z)$ 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, $\{s^{(k)}_n\}$ is zero-mean i.i.d.

S2) $\{s^{(0)}_n\}, \ldots, \{s^{(K)}_n\}$ are jointly statistically independent.

S3) For all k, $\mathrm{E}\{|s^{(k)}_n|^2\} = \sigma_s^2$.

S4) $\mathcal{K}(s^{(0)}_n) \neq 0$.

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

2.2 The Unbiased 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 $s^{(0)}_{n-\nu}$. Using S1)–S3), we may rewrite (1) in terms of the global response $\boldsymbol{q}$ as follows: $J_{\mathrm{m},\nu}(\boldsymbol{q}) = \|\boldsymbol{q} - \mathbf{e}^{(0)}_\nu\|_2^2\, \sigma_s^2$. Denoting MMSE quantities by the subscript "m," it can be shown that in the FIR case S1)–S3) imply that the MMSE channel-plus-estimator is given by $\boldsymbol{q}_{\mathrm{m},\nu} = \boldsymbol{\mathcal{H}}^t(\boldsymbol{\mathcal{H}}^*\boldsymbol{\mathcal{H}}^t)^\dagger\boldsymbol{\mathcal{H}}^*\mathbf{e}_\nu^{(0)}$. A similar expression can be derived for the IIR case (see [2]).

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 and normalize by the receiver gain $q_\nu^{(0)}$. Given that the estimate y_n can be decomposed into signal and interference terms as $y_n = q_\nu^{(0)} s_{n-\nu}^{(0)} + \bar{\boldsymbol{q}}^t \bar{\boldsymbol{s}}(n)$, 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

$$J_{\mathrm{u},\nu}(y_n) \;:=\; \mathrm{E}\big\{|y_n/q_\nu^{(0)} - s_{n-\nu}^{(0)}|^2\big\}. \tag{5}$$

Substituting the decomposition of y_n into (5), and using S1)–S3),

$$J_{\mathrm{u},\nu}(\boldsymbol{q}) \;=\; \frac{\mathrm{E}\big\{|\bar{\boldsymbol{q}}^t\bar{\boldsymbol{s}}(n)|^2\big\}}{|q_\nu^{(0)}|^2} \;=\; \frac{\|\bar{\boldsymbol{q}}\|_2^2}{|q_\nu^{(0)}|^2}\,\sigma_s^2. \tag{6}$$

3 SW Performance

In this section we derive bounds for the UMSE of SW symbol estimators. Section 3.1 outlines our approach, Section 3.2 presents the main results, and Section 3.3 comments on these results. Proofs are not included due to lack of space, but can be found in [2].

3.1 The SW-UMSE Bounding Strategy

Since $y_n = \boldsymbol{q}^t\boldsymbol{s}(n)$ for $\boldsymbol{q} \in \mathcal{Q}_{\mathrm{a}}$, source assumptions S1)-S5) imply that [3]

$$\mathcal{K}(y_n) \;=\; \sum_k \|\boldsymbol{q}^{(k)}\|_4^4\,\mathcal{K}_s^{(k)}, \qquad \sigma_y^2 \;=\; \|\boldsymbol{q}^{(k)}\|_2^2\sigma_s^2, \tag{7}$$

using the shorthand $\mathcal{K}_s^{(k)} = \mathcal{K}(s_n^{(k)})$. This allows a rewrite of the SW criterion (2):

$$\max_{\boldsymbol{q}\in\mathcal{Q}_{\mathrm{a}}\cap\mathcal{Q}_{\mathrm{s}}} \Big|\sum_k \|\boldsymbol{q}^{(k)}\|_4^4\,\mathcal{K}_s^{(k)}\Big|$$

where $\mathcal{Q}_{\mathrm{s}}$ denotes the set of unit-norm global responses: $\mathcal{Q}_{\mathrm{s}} := \{\boldsymbol{q} \text{ s.t. } \|\boldsymbol{q}\|_2 = 1\}$.

Though the SW criterion admits multiple solutions, we are only interested in those that correspond to the estimation of the 0^{th} user's symbols at delay ν. We define the set of global responses *associated*[1] with the {user, delay} pair $\{0,\nu\}$ as follows:

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

The set of SW global responses associated with the $\{0,\nu\}$ pair is then defined by the following local maxima:

$$\{\boldsymbol{q}_{\mathrm{sw},\nu}\} \;:=\; \Big\{\arg\max_{\boldsymbol{q}\in\mathcal{Q}_{\mathrm{a}}\cap\mathcal{Q}_{\mathrm{s}}} \Big|\sum_k \|\boldsymbol{q}^{(k)}\|_4^4\,\mathcal{K}_s^{(k)}\Big|\Big\} \cap \mathcal{Q}_\nu^{(0)}.$$

[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.

It is not possible to write general closed-form expressions for $\{\boldsymbol{q}_{\text{sw},\nu}\}$, making it difficult to characterize their performance.

Consider a reference global response $\boldsymbol{q}_{\text{r},\nu} \in \mathcal{Q}_{\text{a}} \cap \mathcal{Q}_{\text{s}} \cap \mathcal{Q}_{\nu}^{(0)}$. In other words, $\boldsymbol{q}_{\text{r},\nu}$ is an attainable unit-norm response associated with user/delay $\{0, \nu\}$. When $\boldsymbol{q}_{\text{r},\nu}$ is in the vicinity of a $\boldsymbol{q}_{\text{sw},\nu}$ (the meaning of which will be made more precise later), we know that

$$\Big|\sum_k \|\boldsymbol{q}_{\text{sw},\nu}^{(k)}\|_4^4 \, \mathcal{K}_s^{(k)}\Big| \;\geq\; \Big|\sum_k \|\boldsymbol{q}_{\text{r},\nu}^{(k)}\|_4^4 \, \mathcal{K}_s^{(k)}\Big| \;=\; \big|\mathcal{K}(y_{\text{r}})\big|.$$

Thus this $\boldsymbol{q}_{\text{sw},\nu}$ lies in the following set of global responses:

$$\mathcal{Q}_{\text{sw}}(\boldsymbol{q}_{\text{r},\nu}) \;:=\; \Big\{\boldsymbol{q} \;\text{ s.t. }\; \Big|\sum_k \|\boldsymbol{q}^{(k)}\|_4^4 \, \mathcal{K}_s^{(k)}\Big| \geq \big|\mathcal{K}(y_{\text{r}})\big|\Big\} \cap \mathcal{Q}_{\nu}^{(0)} \cap \mathcal{Q}_{\text{s}}. \tag{8}$$

from which an SW-UMSE upper bound may be computed:

$$J_{\text{u},\nu}(\boldsymbol{q}_{\text{sw},\nu}) \;\leq\; \max_{\boldsymbol{q} \in \mathcal{Q}_{\text{sw}}(\boldsymbol{q}_{\text{r},\nu})} J_{\text{u},\nu}(\boldsymbol{q}). \tag{9}$$

Note that (9) avoids explicit consideration of the attainability constraints of $\mathcal{Q}_{\text{a}}$; they are implicitly incorporated via reference $\boldsymbol{q}_{\text{r},\nu} \in \mathcal{Q}_{\text{a}}$. Also note that the tightness of the upper bound (9) will depend on the size and shape of $\mathcal{Q}_{\text{sw}}(\boldsymbol{q}_{\text{r},\nu})$, motivating careful choice of $\boldsymbol{q}_{\text{r},\nu}$. In the sequel we choose the scaled MMSE reference $\boldsymbol{q}_{\text{r},\nu} = \boldsymbol{q}_{\text{m},\nu}/\|\boldsymbol{q}_{\text{m},\nu}\|_2$ (when $\boldsymbol{q}_{\text{m},\nu} \in \mathcal{Q}_{\nu}^{(0)}$) since it is an established benchmark with a closed-form expression.

Two simplifications will ease the evaluation of bound (9). The first is the removal of absolute value signs in the definition (8). Recognize that for $\boldsymbol{q}$ sufficiently close to $\boldsymbol{e}_{\nu}^{(0)}$, $\text{sgn}\big(\sum_k \|\boldsymbol{q}^{(k)}\|_4^4 \, \mathcal{K}_s^{(k)}\big) = \text{sgn}\big(\mathcal{K}_s^{(k)}\big)$, in which case

$$\Big|\sum_k \|\boldsymbol{q}^{(k)}\|_4^4 \, \mathcal{K}_s^{(k)}\Big| \;=\; \text{sgn}\big(\mathcal{K}_s^{(k)}\big) \sum_k \|\boldsymbol{q}^{(k)}\|_4^4 \, \mathcal{K}_s^{(k)}. \tag{10}$$

Our bounds will impose conditions that ensure this behavior.

Next, since both the SW and UMSE criteria 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 the our attention to the set of "de-rotated" global responses $\{\boldsymbol{q} \text{ s.t. } q_{\nu}^{(0)} \in \mathbb{R}^+\}$. For de-rotated responses $\boldsymbol{q} \in \mathcal{Q}_{\text{s}} \cap \mathcal{Q}_{\nu}^{(0)}$, we know $q_{\nu}^{(0)} = \sqrt{1 - \|\bar{\boldsymbol{q}}\|_2^2}$, which implies that such $\boldsymbol{q}$ are completely described by their interference response $\bar{\boldsymbol{q}}$ (as described in Section 2.2). Moreover, these interference responses lie within $\bar{\mathcal{Q}}_{\nu}^{(0)}$, the projection of $\mathcal{Q}_{\nu}^{(0)} \cap \mathcal{Q}_{\text{s}}$ onto $\{\bar{\boldsymbol{q}}\}$:

$$\bar{\mathcal{Q}}_{\nu}^{(0)} \;:=\; \Big\{\bar{\boldsymbol{q}} \text{ s.t. } \sqrt{1 - \|\bar{\boldsymbol{q}}\|_2^2} > \max_{(k,\delta) \neq (0,\nu)} |q_{\delta}^{(k)}|\Big\}.$$

(See Fig. 2 for the construction of $\bar{\mathcal{Q}}_{\nu}^{(0)}$, whose boundary is illustrated by the thick shaded curves.) Using this parameterization, (6) and (7) imply

$$\begin{aligned}
J_{\text{u},\nu}(\boldsymbol{q}_{\text{sw},\nu})/\sigma_s^2 \Big|_{\boldsymbol{q} \in \mathcal{Q}_{\text{s}} \cap \mathcal{Q}_{\nu}^{(0)}} &= \|\bar{\boldsymbol{q}}\|_2^2 / \big(1 - \|\bar{\boldsymbol{q}}\|_2^2\big) \\
\sum_k \|\boldsymbol{q}^{(k)}\|_4^4 \, \mathcal{K}_s^{(k)} \Big|_{\boldsymbol{q} \in \mathcal{Q}_{\text{s}} \cap \mathcal{Q}_{\nu}^{(0)}} &= \big(1 - \|\bar{\boldsymbol{q}}\|_2^2\big)^2 \, \mathcal{K}_s^{(0)} + \sum_k \|\bar{\boldsymbol{q}}^{(k)}\|_4^4 \, \mathcal{K}_s^{(k)}.
\end{aligned} \tag{11}$$

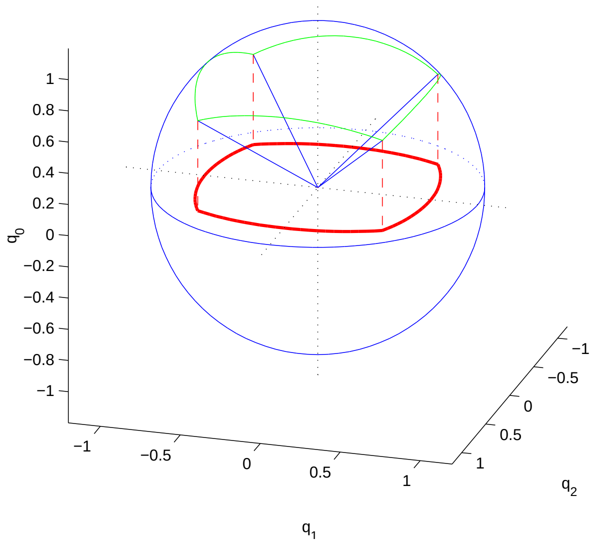

Figure 2: $\bar{\mathcal{Q}}_\nu^{(0)}$, created by projecting $\mathcal{Q}_\nu^{(0)} \cap \mathcal{Q}_\mathrm{s}$ onto the interference space. The boundary of $\bar{\mathcal{Q}}_\nu^{(0)}$ is demarcated by the thick shaded curves.

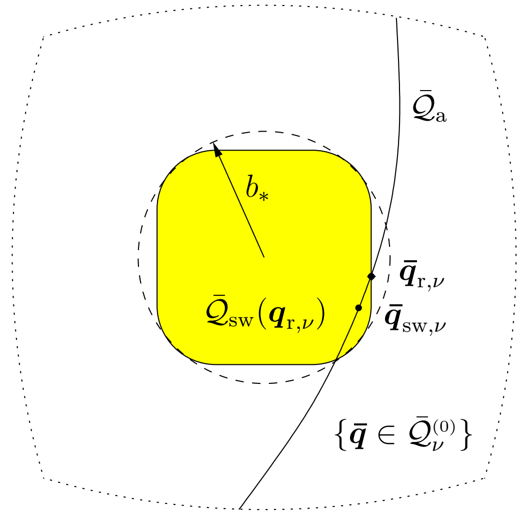

Figure 3: Illustration of SW-UMSE bounding technique in the interference space $\{\bar{\boldsymbol{q}}\}$, i.e., the horizontal plane in Fig. 2.

With the two simplifications above, (9) becomes

$$J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{sw},\nu}) \;\leq\; \max_{\bar{\boldsymbol{q}} \in \bar{\mathcal{Q}}_\mathrm{sw}(\boldsymbol{q}_{\mathrm{r},\nu})} J_{\mathrm{u},\nu}(\bar{\boldsymbol{q}}).$$

where $\bar{\mathcal{Q}}_\mathrm{sw}$ is the following $\{\bar{\boldsymbol{q}}\}$-space projection of $\mathcal{Q}_\mathrm{sw}$:

$$\bar{\mathcal{Q}}_\mathrm{sw}(\boldsymbol{q}_{\mathrm{r},\nu}) := \begin{cases} \left\{ \bar{\boldsymbol{q}} \in \bar{\mathcal{Q}}_\nu^{(0)} \text{ s.t. } \left(1 - \|\bar{\boldsymbol{q}}\|_2^2\right)^2 \mathcal{K}_s^{(0)} + \sum_k \|\bar{\boldsymbol{q}}^{(k)}\|_4^4 \mathcal{K}_s^{(k)} \geq \mathcal{K}(y_\mathrm{r}) \right\}, & \mathcal{K}_s^{(0)} > 0, \\ \left\{ \bar{\boldsymbol{q}} \in \bar{\mathcal{Q}}_\nu^{(0)} \text{ s.t. } \left(1 - \|\bar{\boldsymbol{q}}\|_2^2\right)^2 \mathcal{K}_s^{(0)} + \sum_k \|\bar{\boldsymbol{q}}^{(k)}\|_4^4 \mathcal{K}_s^{(k)} \leq \mathcal{K}(y_\mathrm{r}) \right\}, & \mathcal{K}_s^{(0)} < 0. \end{cases} \tag{12}$$

Finally, since $J_{\mathrm{u},\nu}(\bar{\boldsymbol{q}})$ is strictly increasing in $\|\bar{\boldsymbol{q}}\|_2$ (over its valid range), we claim

$$J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{sw},\nu}) \;\leq\; b_*^2/(1-b_*^2) \quad \text{where} \quad b_* \;:=\; \max_{\bar{\boldsymbol{q}} \in \bar{\mathcal{Q}}_\mathrm{sw}(\boldsymbol{q}_{\mathrm{r},\nu})} \|\bar{\boldsymbol{q}}\|_2. \tag{13}$$

The constrained maximization of b_* can be restated as the following minimization.

$$b_* \;=\; \min_b \text{ s.t. } \left\{ \bar{\boldsymbol{q}} \in \bar{\mathcal{Q}}_\mathrm{sw}(\boldsymbol{q}_{\mathrm{r},\nu}) \Rightarrow \|\bar{\boldsymbol{q}}\|_2 \leq b \right\} \tag{14}$$

Fig. 3 presents a summary of the bounding procedure in the interference response space $\{\bar{\boldsymbol{q}}\}$. The set of attainable interference responses is denoted by $\bar{\mathcal{Q}}_\mathrm{a}$, which can be interpreted as a projection of $\mathcal{Q}_\mathrm{a} \cap \mathcal{Q}_\mathrm{s} \cap \mathcal{Q}_\nu^{(0)}$ onto $\{\bar{\boldsymbol{q}}\}$. Notice that the reference response $\bar{\boldsymbol{q}}_{\mathrm{r},\nu}$ and the SW response $\bar{\boldsymbol{q}}_{\mathrm{sw},\nu}$ both lie on $\bar{\mathcal{Q}}_\mathrm{a}$. Though the exact location of $\bar{\boldsymbol{q}}_{\mathrm{sw},\nu}$ is unknown, we know that it is contained by $\bar{\mathcal{Q}}_\mathrm{sw}(\boldsymbol{q}_{\mathrm{r},\nu})$, depicted in Fig. 3 by the shaded region. Thus, an upper bound on the UMSE of the SW estimator can be calculated using b_*, the maximum interference radius over $\bar{\mathcal{Q}}_\mathrm{sw}(\boldsymbol{q}_{\mathrm{r},\nu})$. As a cautionary note, there exist situations where the shape of $\bar{\mathcal{Q}}_\mathrm{sw}(\boldsymbol{q}_{\mathrm{r},\nu})$ prevents containment by a $\bar{\boldsymbol{q}}$-space ball. In the next section we present conditions (derived in [2]) which avoid these problematic situations.

3.2 The SW-UMSE Bounds

In this section we present SW-UMSE bounds based on the method described in Section 3.1. Proofs appear in [2].

The following kurtosis-based quantities will all prove useful in the sequel:

$$\mathcal{K}_s^{\min} := \min_{0\le k\le K} \mathcal{K}_s^{(k)}, \qquad \mathcal{K}_s^{\max} := \max_{0\le k\le K} \mathcal{K}_s^{(k)}$$
$$\rho_{\min} := \frac{\mathcal{K}_s^{\min}}{\mathcal{K}_s^{(0)}}, \qquad \rho_{\max} := \frac{\mathcal{K}_s^{\max}}{\mathcal{K}_s^{(0)}}$$

Theorem 1. *When $\mathcal{K}(y_{\mathrm{m}})$, the kurtosis of estimates generated by the Wiener estimator associated with the desired user at delay ν, obeys*

$$\begin{cases} \mathcal{K}_s^{(0)} \ \ge\ \mathcal{K}(y_{\mathrm{m}}) \ >\ \big(\mathcal{K}_s^{(0)} + \mathcal{K}_s^{\max}\big)/4, & \textit{for } \mathcal{K}_s^{(0)} > 0, \\ \mathcal{K}_s^{(0)} \ \le\ \mathcal{K}(y_{\mathrm{m}}) \ <\ \big(\mathcal{K}_s^{(0)} + \mathcal{K}_s^{\min}\big)/4, & \textit{for } \mathcal{K}_s^{(0)} < 0, \end{cases} \tag{15}$$

the UMSE of SW estimators associated with the same user/delay can be upper bounded by $J_{\mathrm{u},\nu}\big|_{\mathrm{sw},\nu}^{\max,\mathcal{K}(y_{\mathrm{m}})}$, where

$$J_{\mathrm{u},\nu}\Big|_{\mathrm{sw},\nu}^{\max,\mathcal{K}(y_{\mathrm{m}})} := \begin{cases} \dfrac{1-\sqrt{(\rho_{\max}+1)\frac{\mathcal{K}(y_{\mathrm{m}})}{\mathcal{K}_s^{(0)}}-\rho_{\max}}}{\rho_{\max}+\sqrt{(\rho_{\max}+1)\frac{\mathcal{K}(y_{\mathrm{m}})}{\mathcal{K}_s^{(0)}}-\rho_{\max}}}\,\sigma_s^2, & \textit{for } \mathcal{K}_s^{(0)} > 0, \\[2ex] \dfrac{1-\sqrt{(\rho_{\min}+1)\frac{\mathcal{K}(y_{\mathrm{m}})}{\mathcal{K}_s^{(0)}}-\rho_{\min}}}{\rho_{\min}+\sqrt{(\rho_{\min}+1)\frac{\mathcal{K}(y_{\mathrm{m}})}{\mathcal{K}_s^{(0)}}-\rho_{\min}}}\,\sigma_s^2, & \textit{for } \mathcal{K}_s^{(0)} < 0. \end{cases} \tag{16}$$

Furthermore, (15) guarantees the existence of a SW estimator associated with this user/delay when $\boldsymbol{q}$ is FIR.

While Theorem 1 presents a closed-form SW-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 $J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu}) < J_{\mathrm{o}}\sigma_s^2$, where*

$$J_{\mathrm{o}} := \begin{cases} 2\sqrt{(1+\rho_{\max})^{-1}}-1 & \mathcal{K}_s^{(0)}>0,\ \mathcal{K}_s^{\min}\ge 0 \\ \frac{1-\sqrt{1-(3-\rho_{\max})(1+\rho_{\min})/4}}{\rho_{\min}+\sqrt{1-(3-\rho_{\max})(1+\rho_{\min})/4}}, & \mathcal{K}_s^{(0)}>0,\ \mathcal{K}_s^{\min}<0,\ \mathcal{K}_s^{\min}\ne-\mathcal{K}_s^{(0)} \\ \frac{3-\rho_{\max}}{5+\rho_{\max}} & \mathcal{K}_s^{(0)}>0,\ \mathcal{K}_s^{\min}<0,\ \mathcal{K}_s^{\min}=-\mathcal{K}_s^{(0)} \\ 2\sqrt{(1+\rho_{\min})^{-1}}-1 & \mathcal{K}_s^{(0)}<0,\ \mathcal{K}_s^{\max}\le 0 \\ \frac{1-\sqrt{1-(3-\rho_{\min})(1+\rho_{\max})/4}}{\rho_{\max}+\sqrt{1-(3-\rho_{\min})(1+\rho_{\max})/4}}, & \mathcal{K}_s^{(0)}<0,\ \mathcal{K}_s^{\max}>0,\ \mathcal{K}_s^{\max}\ne-\mathcal{K}_s^{(0)} \\ \frac{3-\rho_{\min}}{5+\rho_{\min}} & \mathcal{K}_s^{(0)}<0,\ \mathcal{K}_s^{\max}>0,\ \mathcal{K}_s^{\max}=-\mathcal{K}_s^{(0)} \end{cases} \tag{17}$$

the UMSE of SW estimators associated with the same user/delay can be bounded as follows:

$$J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu}) \ \le\ J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{sw},\nu}) \ \le\ J_{\mathrm{u},\nu}\Big|_{\mathrm{sw},\nu}^{\max,\mathcal{K}(y_{\mathrm{m}})} \ \le\ J_{\mathrm{u},\nu}\Big|_{\mathrm{sw},\nu}^{\max,J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})},$$

where

$$J_{\mathrm{u},\nu}\Big|_{\mathrm{sw},\nu}^{\max,J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})} := \tag{18}$$

$$\begin{cases} \dfrac{1-\sqrt{(1+\rho_{\max})\left(1+\frac{J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})}{\sigma_s^2}\right)^{-2}-\rho_{\max}}}{\rho_{\max}+\sqrt{(1+\rho_{\max})\left(1+\frac{J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})}{\sigma_s^2}\right)^{-2}-\rho_{\max}}}\,\sigma_s^2 & \mathcal{K}_s^{(0)}>0,\ \ \mathcal{K}_s^{\min}\geq 0 \\[2ex] \dfrac{1-\sqrt{(1+\rho_{\max})\left(1+\frac{J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})}{\sigma_s^2}\right)^{-2}\left(1+\rho_{\min}\frac{J^2_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})}{\sigma_s^4}\right)-\rho_{\max}}}{\rho_{\max}+\sqrt{(1+\rho_{\max})\left(1+\frac{J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})}{\sigma_s^2}\right)^{-2}\left(1+\rho_{\min}\frac{J^2_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})}{\sigma_s^4}\right)-\rho_{\max}}}\,\sigma_s^2 & \mathcal{K}_s^{(0)}>0,\ \ \mathcal{K}_s^{\min}<0 \\[2ex] \dfrac{1-\sqrt{(1+\rho_{\min})\left(1+\frac{J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})}{\sigma_s^2}\right)^{-2}-\rho_{\min}}}{\rho_{\min}+\sqrt{(1+\rho_{\min})\left(1+\frac{J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})}{\sigma_s^2}\right)^{-2}-\rho_{\min}}}\,\sigma_s^2 & \mathcal{K}_s^{(0)}<0,\ \ \mathcal{K}_s^{\max}\leq 0 \\[2ex] \dfrac{1-\sqrt{(1+\rho_{\min})\left(1+\frac{J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})}{\sigma_s^2}\right)^{-2}\left(1+\rho_{\max}\frac{J^2_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})}{\sigma_s^4}\right)-\rho_{\min}}}{\rho_{\min}+\sqrt{(1+\rho_{\min})\left(1+\frac{J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})}{\sigma_s^2}\right)^{-2}\left(1+\rho_{\max}\frac{J^2_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})}{\sigma_s^4}\right)-\rho_{\min}}}\,\sigma_s^2 & \mathcal{K}_s^{(0)}<0,\ \ \mathcal{K}_s^{\max}>0 \end{cases}$$

Furthermore, (17) guarantees the existence of a SW estimator associated with this user/delay when $\boldsymbol{q}$ is FIR.

Equation (18) leads to an elegant approximation of the *extra* UMSE of SW estimators:

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

Theorem 3. *If $J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu}) < J_{\mathrm{o}}\sigma_s^2$, then the extra UMSE of SW estimators can be bounded as $\mathcal{E}_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{sw},\nu}) \leq \mathcal{E}_{\mathrm{u},\nu}\big|_{\mathrm{c},\nu}^{\max,J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})}$, where*

$$\begin{aligned} \mathcal{E}_{\mathrm{u},\nu}\Big|_{\mathrm{c},\nu}^{\max,J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})} &:= J_{\mathrm{u},\nu}\Big|_{\mathrm{sw},\nu}^{\max,J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})} - J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu}) \\ &= \begin{cases} \frac{1}{2\sigma_s^2}\rho_{\max}J^2_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu}) + \mathcal{O}\big(J^3_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})\big) & \mathcal{K}_s^{(0)}>0,\ \ \mathcal{K}_s^{\max}\geq 0 \\ \frac{1}{2\sigma_s^2}(\rho_{\max}-\rho_{\min})J^2_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu}) + \mathcal{O}\big(J^3_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})\big) & \mathcal{K}_s^{(0)}>0,\ \ \mathcal{K}_s^{\max}< 0 \\ \frac{1}{2\sigma_s^2}\rho_{\min}J^2_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu}) + \mathcal{O}\big(J^3_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})\big) & \mathcal{K}_s^{(0)}<0,\ \ \mathcal{K}_s^{\max}\leq 0 \\ \frac{1}{2\sigma_s^2}(\rho_{\min}-\rho_{\max})J^2_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu}) + \mathcal{O}\big(J^3_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})\big) & \mathcal{K}_s^{(0)}<0,\ \ \mathcal{K}_s^{\max}> 0 \end{cases} \end{aligned} \tag{19}$$

Equation (19) implies that the extra UMSE of SW estimators is upper bounded by approximately the *square* of the minimum UMSE. Fig. 4 plots the upper bound on SW-UMSE and extra SW-UMSE from (18) as a function of $J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})/\sigma_s^2$ for various values of $\rho_{\min}$ and $\rho_{\max}$. The second-order approximation based on (19) appears very good for all but the largest values of UMSE.

3.3 Comment: Implicit Incorporation of $\mathcal{Q}_{\mathrm{a}}$

First, recall that the SW-UMSE bounding procedure incorporated $\mathcal{Q}_{\mathrm{a}}$, the set of attainable global responses, *only* in the requirement that $\boldsymbol{q}_{\mathrm{r},\nu} \in \mathcal{Q}_{\mathrm{a}} \cap \mathcal{Q}_{\mathrm{s}} \cap \mathcal{Q}_{\nu}^{(0)}$. Thus

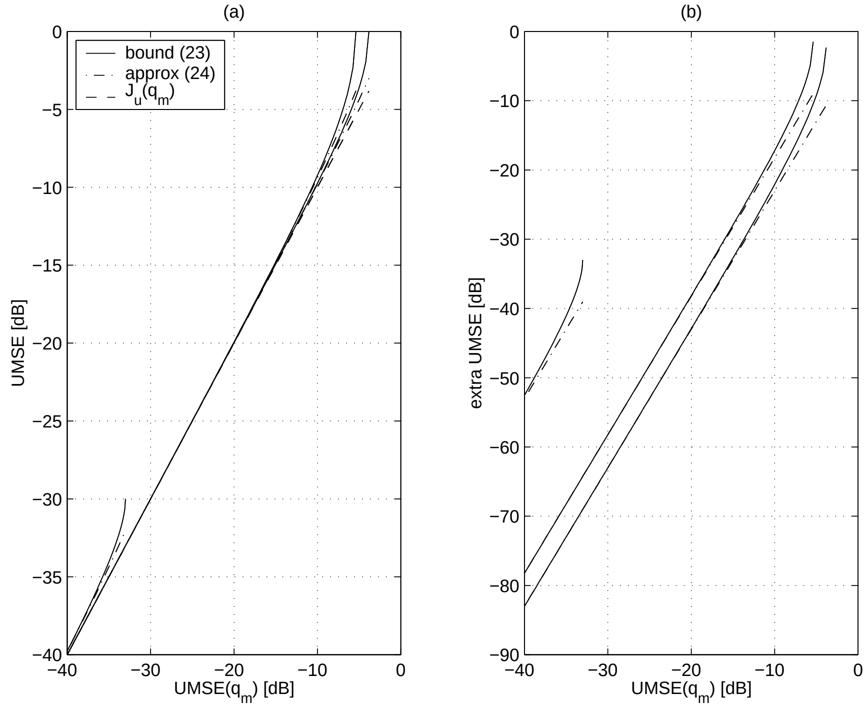

Figure 4: Upper bound on (a) SW-UMSE and (b) extra SW-UMSE versus $J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})$ (when $\sigma_s^2 = 1$) from (18) with second-order approximation from (19). From left to right, $\{\rho_{\min}, \rho_{\max}\} = \{1000, 0\}$, $\{1, -2\}$, and $\{1, 0\}$.

Theorems 1–3, written under the reference choice $\boldsymbol{q}_{\mathrm{r},\nu} = \boldsymbol{q}_{\mathrm{m},\nu}/\|\boldsymbol{q}_{\mathrm{m},\nu}\|_2 \in \mathcal{Q}_{\mathrm{a}} \cap \mathcal{Q}_{\mathrm{s}} \cap \mathcal{Q}_\nu^{(0)}$, implicitly incorporate the channel and/or estimator constraints that define $\mathcal{Q}_{\mathrm{a}}$. For example, if $\boldsymbol{q}_{\mathrm{m},\nu}$ is the MMSE response constrained to the set of causal IIR estimators, then SW-UMSE bounds based on this $\boldsymbol{q}_{\mathrm{m},\nu}$ will implicitly incorporate the causality constraint. The implicit incorporation of the attainable set $\mathcal{Q}_{\mathrm{a}}$ makes these bounding theorems quite general and easy to use.

4 Numerical Examples

Here we present the results of an experiment which compares our UMSE bounds to the UMSE characterizing SW estimators found by gradient descent.[2] Ten super-Gaussian sources ($\mathcal{K}_s^{(k)} = 2$) were mixed using matrix $\boldsymbol{\mathcal{H}}$ with real-valued zero-mean Gaussian entries. The estimator observed the mixture in the presence of AWGN (at -20dB) and generated estimates of a particular source using 8 adjustable parameters. Note that the number of sensors is less than the number of sources and that noise is present, implying that $\boldsymbol{\mathcal{H}}$ is not full column rank and perfect estimation is not possible.

Fig. 5(a) plots the UMSE upper bounds $J_{\mathrm{u},\nu}\big|_{\mathrm{sw},\nu}^{\max,\mathcal{K}(y_{\mathrm{m}})}$ and $J_{\mathrm{u},\nu}\big|_{\mathrm{sw},\nu}^{\max,J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})}$ for comparison with $J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{sw},\nu})$. As a means of "zooming in" on the small differences in UMSE, Fig. 5(b) plots the extra-UMSE upper bounds $\mathcal{E}_{\mathrm{u},\nu}\big|_{\mathrm{c},\nu}^{\max,\mathcal{K}(\boldsymbol{q}_{\mathrm{m},\nu})}$ and $\mathcal{E}_{\mathrm{u},\nu}\big|_{\mathrm{c},\nu}^{\max,J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})}$. The $J_{\mathrm{u},\nu}(\boldsymbol{q}_{\mathrm{m},\nu})$-based bounds are denoted by solid lines, the $\mathcal{K}(\boldsymbol{q}_{\mathrm{m},\nu})$-based bounds are denoted

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

by •'s, and the gradient-descent values are denoted by ×'s. Note the tightness of the bounds for all but the largest values of $J_{u,\nu}(\boldsymbol{q}_{m,\nu})$.

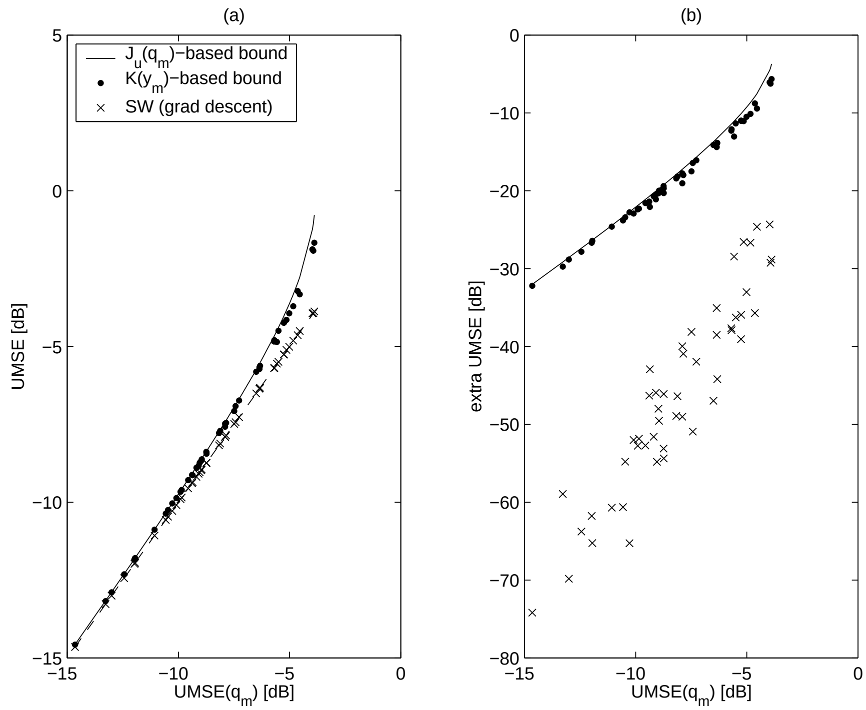

Figure 5: Bounds on SW-UMSE for $N_f = 8$, 10 sources with $\mathcal{K}_s^{(k)} = 2$, AWGN at -20dB, and random $\boldsymbol{\mathcal{H}}$.

5 Conclusions

In this paper we have derived conditions under which SW estimators exist and derived bounds for the UMSE of SW estimators. The existence conditions are simple tests which guarantee a SW estimator for the desired user at a particular delay, and these existence arguments have been proven for vector-valued FIR channels and constrained vector-valued FIR estimators. The first bound is a function of the kurtosis of the MMSE estimates, while the second bound is a function of the minimum UMSE. Analysis of the second bound shows that the extra UMSE of SW estimators is upper bounded by approximately the square of the minimum UMSE. Thus, SW estimators are very close (in a MSE sense) to optimum linear estimators when the minimum MSE is small. Numerical simulations suggest that the bounds are reasonably tight.

References

[1] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Trans. on Info. Theory*, vol. 36, no. 2, pp. 312-21, Mar. 1990.

[2] P. Schniter, "Blind Estimation Without Priors: Performance, Convergence, and Efficient Implementation," *Ph.D. dissertation*, Cornell University, Ithaca, NY, May 2000.

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