An Shalvi-Weinstein Initialization Lemma

Phil Schniter

February 12, 1999

1 Introduction

In this report we present a simple sufficient condition under which small-stepsize Shalvi-Weinstein (SW) [1] is guaranteed to converge to a *particular* minimum. The result is a weaker form of a kurtosis-based sufficient condition presented by Li and Ding in [2] and involves very simple SIR-based requirements, making it convenient for *local* convergence analysis of SW. For simplicity, we assume an i.i.d. source and real-valued quantities in the absence of noise, though we do not place any restrictions on the rank of the channel matrix $\mathbf{H}$.

The following model is assumed. $\mathbf{H}$ denotes the linear channel taking the time-n source vector $\mathbf{s}(n)$ to the time-n received vector $\mathbf{r}(n)$:

$$\mathbf{r}(n) = \mathbf{H}\mathbf{s}(n). \tag{1}$$

From the received vector, the equalizer $\mathbf{f}$ generates a linear estimate of the δ^{th} element of the source vector at time n:

$$\hat{s}_\delta(n) = \mathbf{f}^t\mathbf{r}(n). \tag{2}$$

Most of our results will use the $N \times 1$ "global" or "system" response vector $\mathbf{q}$ defined

$$\mathbf{q} \stackrel{\triangle}{=} \mathbf{H}^t\mathbf{f} = (q_0, \dots, q_\delta, \dots, q_{N-1})^t, \tag{3}$$

as well as the $(N-1) \times 1$ "interference response" $\bar{\mathbf{q}}$:

$$\bar{\mathbf{q}} \stackrel{\triangle}{=} (q_0, \dots, q_{\delta-1}, q_{\delta+1}, \dots, q_{N-1})^t, \tag{4}$$

The mean-squared estimation error (MSE) associated with estimation of the δ^{th} element in $\mathbf{s}(n)$ using system response $\mathbf{q}$ will be denoted

$$\mathrm{MSE}_\delta(\mathbf{q}) \stackrel{\triangle}{=} \mathrm{E}\big\{\big(s_\delta(n) - \mathbf{q}^t\mathbf{s}(n)\big)^2\big\}. \tag{5}$$

As a last point, note that we will refer to the index of the largest value of a given vector using the term "cursor."

2 Background

For background, we paraphrase Li and Dings' Theorem (originally appearing as Theorem 4.1 in [2]).

Theorem 1 (Li & Ding). *Consider an initial system response* $\mathbf{q} \in \mathrm{row}(\mathbf{H})$ *with cursor* δ *and i.i.d. source. If*

$$\frac{\|\mathbf{q}\|_4^4}{\|\mathbf{q}\|_2^4} > \frac{1}{2}, \tag{6}$$

then small-stepsize SW adaptation of equalizer $\mathbf{f}$ *will converge to a system response* $\mathbf{q}_{\text{final}} \in \mathrm{row}(\mathbf{H})$ *with cursor* δ*.*

3 The Main Result

The main result appears below. It can be considered as a sufficient condition for (6) whose simple form makes it potentially useful for the analysis of SW initializations.

Lemma 1. *Consider initial system response* $\mathbf{q} \in \mathrm{row}(\mathbf{H})$ *and i.i.d. source. If*

$$\frac{q_\delta^2}{\|\mathbf{q}\|_2^2} > 1 - \tfrac{N-1}{N}\Big(1-\sqrt{1-\tfrac{N}{2N-2}}\Big) \triangleq \beta(N), \tag{7}$$

then small-stepsize SW adaptation of equalizer $\mathbf{f}$ *will converge to a system response* $\mathbf{q}_{\text{final}} \in \mathrm{row}(\mathbf{H})$ *with cursor* δ*.*

Proof. We start by noticing that condition (6) in Theorem 1 is insensitive to any scaling of the system response. Hence we can assume, w.l.o.g., that $\|\mathbf{q}\|_2^2 = 1$. Under this normalization, $q_\delta^2 = 1 - \|\bar{\mathbf{q}}\|_2^2$, allowing us to write $\|\mathbf{q}\|_4^4 = \|\bar{\mathbf{q}}\|_4^4 + (1-\|\bar{\mathbf{q}}\|_2^2)^2$ which allows (6) to be rewritten as follows:

$$\frac{1}{2} < \|\bar{\mathbf{q}}\|_4^4 + (1-\|\bar{\mathbf{q}}\|_2^2)^2. \tag{8}$$

Next, recall that when subject to the constraint $\|\bar{\mathbf{q}}\|_2^2 = b$, $\|\bar{\mathbf{q}}\|_4$ is minimized by a vector $\bar{\mathbf{q}}$ of the form $\bar{\mathbf{q}} = \Big(\pm\sqrt{\tfrac{b}{N-1}}, \pm\sqrt{\tfrac{b}{N-1}}, \ldots, \pm\sqrt{\tfrac{b}{N-1}}\Big)^t \in \mathbb{R}^{N-1}$. It follows that

$$\min_{\|\bar{\mathbf{q}}\|_2^2=b} \|\bar{\mathbf{q}}\|_4^4 = (N-1)\left(\frac{b}{N-1}\right)^2 = \frac{b^2}{N-1}.$$

Thus, when $\|\bar{\mathbf{q}}\|_2^2 = b$, the following condition is sufficient for the satisfaction of (8):

$$\frac{1}{2} < \frac{b^2}{N-1} + (1-b)^2. \tag{9}$$

By solving the quadratic we can obtain a sufficient condition for (9):

$$\|\bar{\mathbf{q}}\|_2^2 = b < \tfrac{N-1}{N}\Big(1-\sqrt{1-\tfrac{N}{2N-2}}\Big).$$

Recalling that $q_\delta^2 = 1 - \|\bar{\mathbf{q}}\|_2^2$, the preceding condition is equivalent to:

$$q_\delta^2 > 1 - \tfrac{N-1}{N}\Big(1-\sqrt{1-\tfrac{N}{2N-2}}\Big). \tag{10}$$

Finally, the scale-invariance of (6) can be restored by dividing the left hand side of (10) by $\|\mathbf{q}\|_2^2$, giving condition (7) stated by this Lemma. □

Figure 1 presents a geometrical view of Lemma 1 for the case $N = 3$, while Table 1 lists values of the threshold $\beta(N)$.

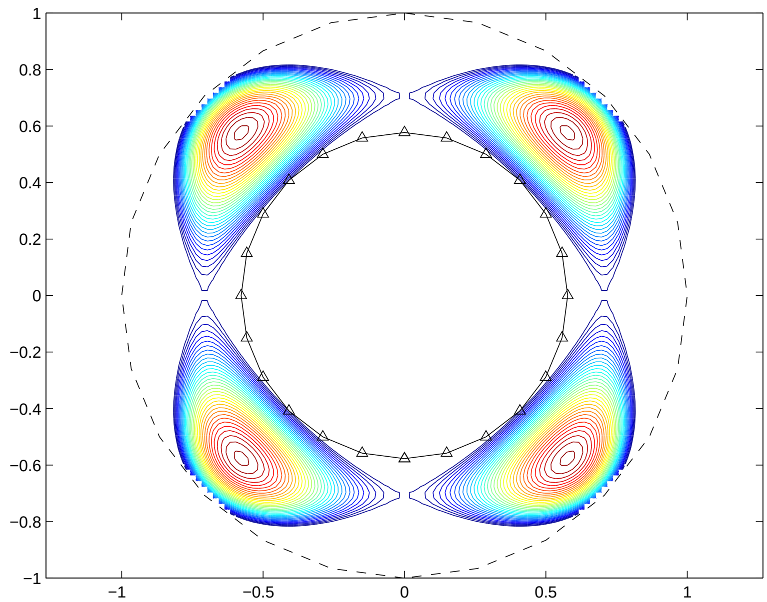

Figure 1: Projection of 3-dimensional $\mathbf{q}$-space unit sphere onto 2-dimensional $\bar{\mathbf{q}}$-space, with contours identifying the regions where Li & Ding condition (6) is not satisfied. Lemma 1 claims that all unit-normalized system responses $\mathbf{q}$ whose interference responses $\bar{\mathbf{q}}$ lie within the triangle-demarcated ring are guaranteed to lie inside the SW region of convergence associated with system delay δ.

Table 1: Examples values for the quantity $\beta(N)$ discussed by Lemma 1

N	2	3	5	10	50	100	1000
$\beta(N)$	0.5	0.667	0.690	0.700	0.706	0.706	0.707

4 A Few Implications

Lemma 1 has a number of implications, a few of which are given below. Corollary 1 presents a simple condition on the channel matrix $\mathbf{H}$ which guarantees that a ν-spike initialization will converge to a system response whose cursor equals that of the ν^{th} row of $\mathbf{H}$. Corollary 2 presents a simple condition on the initial system response under which the equalizer trajectory will be immune to attraction by saddle points. Corollary 3 presents a simple test condition sufficient to guarantee the existence of a Wiener receiver in the ROC of the SW receiver associated with the same "system delay".

Corollary 1 (Single-spike). *Consider an i.i.d. source and an equalizer initialization with a single-spike at index ν, and denote the $(\nu, \delta)^{th}$ element of $\mathbf{H}$ by $h_{\nu,\delta}$ and the ν^{th} row of $\mathbf{H}$ by $\mathbf{h}_\nu^t$. If*

$$\frac{h_{\nu,\delta}^2}{\|\mathbf{h}_\nu\|_2^2} \; > \; 1 - \tfrac{N-1}{N}\left(1-\sqrt{1-\tfrac{N}{2N-2}}\right),$$

then small-stepsize SW adaptation of equalizer $\mathbf{f}$ *will converge to a system response* $\mathbf{q}_{\text{final}} \in \text{row}(\mathbf{H})$ *with cursor* δ.

Proof. This is a simple consequence of the fact that a single-spike equalizer initialization creates an initial system response equal to a particular row of $\mathbf{H}$. □

Corollary 2 (Saddles). *Consider an i.i.d. source and an initial system response* $\mathbf{q} \in \text{row}(\mathbf{H})$. *If*

$$\frac{q_\delta^2}{\|\mathbf{q}\|_2^2} \; > \; 1 - \tfrac{N-1}{N}\left(1-\sqrt{1-\tfrac{N}{2N-2}}\right), \tag{11}$$

then the equalizer trajectory associated with small-stepsize SW adaptation of equalizer $\mathbf{f}$ *will avoid capture by saddle points.*

Proof. As with the CM criterion, SW saddle points have system responses characterized by elements which are either zero-valued or of equal magnitude [3]. If we evaluate the quantity $\frac{\|\mathbf{q}\|_4^4}{\|\mathbf{q}\|_2^4}$ over all possible saddle points $\mathbf{q}$, we find that it reaches a minimum value of 1/2, meaning that all saddles lie outside of the convergence region identified in (6). Thus, since any small-stepsize SW trajectory whose initial point obeys (6) must consist only of other points satisfying (6), such a trajectory would avoid all saddle points. Finally, satisfaction of (11) is sufficient for satisfaction of (6).

□

Corollary 3 (Wiener in ROC). *If there exists* $\mathbf{q} \in \text{row}(\mathbf{H})$ *such that*

$$\frac{\text{MSE}_\delta(\mathbf{q})}{\sigma_s^2} < \frac{N-1}{N}\left(1-\sqrt{1-\frac{N}{2N-2}}\right).$$

then, assuming an i.i.d. source, the Wiener receiver $\mathbf{q}_m$ *with cursor* δ *lies within the region-of-convergence associated with the SW receivers of cursor* δ.

Proof. This follows from the fact that

$$\frac{\text{MSE}_\delta(\mathbf{q})}{\sigma_s^2} > \frac{\text{MMSE}_\delta}{\sigma_s^2} = 1 - \frac{q_{\text{m}\delta}^2}{\|\mathbf{q}_\text{m}\|_2^2}.$$

□

5 Applications

Lemma 1 enables the analysis of CDMA applications of SW. Specifically, it can be used to make statements about acquisition of a *particular* user under various initialization strategies and channel conditions.

In a more general sense, it motivates SW initialization strategies that use, as a first stage, blind maximization of SIR. Since this can be accomplished via SOS-based techniques, this has the potential of combining the speed of SOS (as a byproduct of "good" initialization) with the robustness of SW (as characterized by the steady-state performance).

6 Extensions

The results would benefit from generalization to the case where there exist a number of different sources with different kurtoses. Under proper construction of $\mathbf{H}$, this should subsume the noisy case.

References

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

[2] Y. Li and Z. Ding, "Convergence analysis of finite length blind adaptive equalizers," *IEEE Trans. Signal Processing*, vol. 43, no. 9, pp. 2120-29, Sep. 1995.

[3] C.R. Johnson, Jr., P. Schniter, I. Fijalkow, L. Tong, et al., J.R. Treichler, "The Core of FSE-CMA Behavior Theory," *Unsupervised Adaptive Learning* (ed. S. Haykin), New York, NY: Wiley, 1999.