UNDERMODELED EQUALIZATION: EXTREMA OF THE GODARD/SHALVI-WEINSTEIN CRITERION

Phillip A. Regalia
Institut National des Télécommunications
9, rue Charles Fourier
F-91011 Evry cedex France
Fax: +33 1 60 76 44 33
regalia@int-evry.fr

Mamadou Mboup
UFR Mathématiques et Informatique
Université René Descartes
45, rue des Saints-Pères
F-75270 Paris cedex 06, France
mboup@math-info.univ-paris5.fr

ABSTRACT

The more promising results to date in blind equalization are restricted to so-called sufficient-order settings, in which all configurations in the combined (channel-equalizer cascade) impulse response space are attainable for some setting of the equalizer coefficients. Here we address issues related to equalizer undermodelling, in which only a proper subset within the combined impulse response space is attainable by adjusting the equalizer coefficients. We derive an analytic characterization of stationary points for the Godard and Shalvi-Weinstein criteria in undermodeled cases, an establish relations to the convergence of super-exponential algorithms in undermodeled cases.

1 INTRODUCTION

Blind adaptation algorithms for channel equalization purposes are an attractive alternative to their training sequence based counterparts, as blind methods obviate the need to send training sequences. The more promising results to date on blind equalization concern fractionally spaced systems: Provided the equivalent channel impulse response is strictly finite in duration, the subchannels have no common zeros, and the equalizer impulse response length is chosen adequately, the more popular blind equalization criteria can claim to be free from local minima traps [1]. Although multiple minima are generically present, each achieves perfect channel equalization.

Here we examine a more realistic *undermodeled* scenario (in which "undermodeled" is defined in the next section), which arises in practice if the equalizer length is chosen insufficient, and/or the subchannels have common zeros, and/or background noise is present, among other harsh realities. The idealized scenario, featuring multiple minima with each achieving perfect channel equalization, is deformed into a more daunting scenario, in which multiple minima are still present, but now offering disparate performance levels.

We present in this paper a characterization of stationary points, in undermodeled cases, for a family of blind equalization criteria. The family includes the popular Godard [2] (or CMA [3]) and Shalvi-Weinstein [4] criteria, as well as the super-exponential [5] algorithms. For ease of notation, we consider real channels and signal constellations in this work; the case of complex signal constellations and complex channels is detailed in [8].

2 PROBLEM SETTING

We review in this section some basic concepts and considerations for the following channel-equalizer cascade:

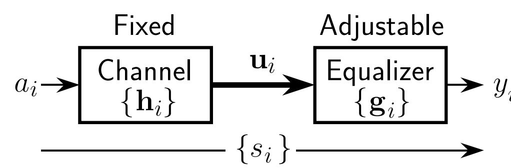

Figure 1: Channel-equalizer cascade

The equalizer input is obtained after demodulation and sampling of the receiver output, giving the N-element vector

$$\mathbf{u}_i = \sum_k \mathbf{h}_k \, a_{i-k}$$

where $\{\mathbf{h}_k\}$ is the single-input–N-output channel impulse response, and $\{a_i\}$ is the source sequence. The baud-rate case corresponds to $N = 1$ and the fractionally-spaced case corresponds to $N \geq 2$. The equalizer output is

$$y_i = \sum_{k=0}^{L} \mathbf{g}_k \, \mathbf{u}_{i-k}$$

in terms of the N-input–single-output impulse response $\{\mathbf{g}_i\}$, of length $L+1$.

The combined response $\{s_k\}$, which maps the source sequence $\{a_i\}$ to the equalizer output $\{y_i\}$, is the convolution of $\{\mathbf{h}_i\}$ and $\{\mathbf{g}_i\}$:

$$s_k = \sum_{i=0}^{L} \mathbf{g}_i \, \mathbf{h}_{k-i}$$

In matrix form this becomes

$$\underbrace{\begin{bmatrix} s_0 \\ s_1 \\ \vdots \\ s_L \\ s_{L+1} \\ \vdots \end{bmatrix}}_{\mathbf{s}} = \underbrace{\begin{bmatrix} \mathbf{h}_0^T & \mathbf{0}^T & \cdots & \mathbf{0}^T \\ \mathbf{h}_1^T & \mathbf{h}_0^T & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}^T \\ \mathbf{h}_L^T & \cdots & \mathbf{h}_1^T & \mathbf{h}_0^T \\ \mathbf{h}_{L+1}^T & \mathbf{h}_L^T & \cdots & \mathbf{h}_1^T \\ \vdots & \ddots & \ddots & \vdots \end{bmatrix}}_{\mathbf{H}} \begin{bmatrix} \mathbf{g}_0^T \\ \mathbf{g}_1^T \\ \vdots \\ \mathbf{g}_L^T \end{bmatrix}$$

It is clear that the vector $\mathbf{s}$ is restricted to the range space of the matrix $\mathbf{H}$; this range space is denoted $\mathcal{S}_A$ as in [1]. We denote by $\mathcal{P}_A$ the orthogonal projection operator onto $\mathcal{S}_A$; this may be written as $\mathbf{H}(\mathbf{H}^T\mathbf{H})^\sharp\mathbf{H}^T$, where superscript $\sharp$ denotes (pseudo-) inversion.

An equalizer $\{\mathbf{g}_i\}$ will be denoted *sufficient order* if $\mathcal{P}_A = \mathbf{I}$ (the identity), and *undermodeled* if $\mathcal{P}_A \neq \mathbf{I}$. In the sufficient order case, an arbitrary configuration in the combined response can be attained for some setting of the equalizer coefficients $\{\mathbf{g}_i\}$, including an ideal response of the form

$$\mathbf{s} = \mathbf{e}_n \triangleq [\cdots, 0, \underbrace{1}_{n}, 0, \cdots]$$

where n is any integer; this combined response represents a pure delay of n samples. The undermodeled case will, in general, preclude the possibility of such an ideal combined response being attainable.

A popular criterion for equalization, originating in Donoho's work [7], involves normalized cumulants of the form $\text{cum}_{2p}(y_i)/[\text{cum}_2(y_i)]^p$, in which $\text{cum}_{2p}(\cdot)$ is the cumulant of order $2p$ of a random variable. If the source $\{a_i\}$ is i.i.d., then

$$\frac{\text{cum}_{2p}(y_i)}{[\text{cum}_2(y_i)]^p} = \underbrace{\left(\sum_k s_k^{2p} \Big/ \Big(\sum_k s_k^2\Big)^p\right)}_{[D_{2p}(\mathbf{s})]^{2p}} \frac{\text{cum}_{2p}(a_i)}{[\text{cum}_2(a_i)]^p}$$

involving the multiplicative factor

$$D_{2p}(\mathbf{s}) = \frac{\|\mathbf{s}\|_{2p}}{\|\mathbf{s}\|_2}$$

which in turn is the ratio of different sequence norms

$$\|\mathbf{s}\|_{2p} = \begin{cases} \Big(\sum_k |s_k|^{2p}\Big)^{1/2p}, & 2p < \infty; \\ \sup_k |s_k|, & p = \infty. \end{cases}$$

Maximizing $|\,\text{cum}_{2p}(y_i)/[\text{cum}_2(y_i)]^p|$, as in the Shalvi-Weinstein algorithm [4], amounts to seeking the maximum of $D_4(\mathbf{s})$ if the source is i.i.d., while the same applies to the Godard algorithm [1] if in addition the source has a negative fourth-order cumulant. The following properties may be shown:

- $D_{2p}(\alpha\mathbf{s}) = D_{2p}(\mathbf{s})$ for any nonzero scalar α;
- $0 < D_{2p}(\mathbf{s}) \leq 1$ for all $\mathbf{s} \neq \mathbf{0}$, and $D_{2p}(\mathbf{s}) = 1$ if and only if $\mathbf{s} = \alpha\mathbf{e}_n$ with n any integer;
- In the sufficient order case ($\mathcal{P}_A = \mathbf{I}$), each local maximum of $D_{2p}(\mathbf{s})$ is of the form $\mathbf{s} = \alpha\mathbf{e}_n$, for any $p \geq 2$.

For the undermodeled case $\mathcal{P}_A \neq \mathbf{I}$, by contrast, the set of maxima varies with p. A characterization of stationary points, obtained under the constraint that $\mathbf{s}$ be restricted to $\mathcal{S}_A$, is the subject of the next section.

3 STATIONARY POINTS OF $D_{2p}(\mathbf{s})$

We begin with some vector notation which will simplify further expressions. Given two vectors $\mathbf{r}$ and $\mathbf{s}$ (attainable or not), their inner product is denoted

$$\langle \mathbf{r}, \mathbf{s} \rangle = \sum_k s_k\, r_k,$$

and their Hadamard (or componentwise) product will be denoted

$$\mathbf{r} \odot \mathbf{s}, \quad \text{with } k^{th} \text{ component } [\mathbf{r} \odot \mathbf{s}]_k = r_k\, s_k.$$

The Hadamard exponent follows similarly as

$$\mathbf{s}^{\odot m} = \underbrace{\mathbf{s} \odot \cdots \odot \mathbf{s}}_{m \text{ terms}}$$

whose k^{th} component is s_k^m.

Theorem 1 *Let $\mathcal{P}_A$ be the orthogonal projection operator onto $\mathcal{S}_A$. A candidate $\mathbf{s} \in \mathcal{S}_A$ is a stationary point of $D_{2p}(\mathbf{s})$ if and only if*

$$\mathcal{P}_A\big(\mathbf{s}^{\odot(2p-1)}\big) = \alpha\mathbf{s}, \qquad \textit{for some scalar } \alpha.$$

If $\mathbf{s}$ is scaled to unit ℓ_2 norm, then $\sqrt[2p]{\alpha} = D_{2p}(\mathbf{s})$, the value obtained at the stationary point.

Proof: If $\mathbf{s}$ and $\mathbf{r}$ are two vectors in $\mathcal{S}_A$, then $\mathbf{q}(t) = \mathbf{s} + t\mathbf{r}$ remains in $\mathcal{S}_A$ for all t. We introduce the moments of orders 2 and $2p$ of $\mathbf{q}(t)$ as

$$m_2(t) = \sum_k |q_k(t)|^2, \quad m_{2p}(t) = \sum_k |q_k(t)|^{2p},$$

and we suppose that $\mathbf{s}$ is scaled to unit ℓ_2 norm for simplicity, giving $m_2(0) = 1$.

Introduce now the functional

$$F_{2p}\big(\mathbf{q}(t)\big) = \Big[D_{2p}\Big(\mathbf{q}(t)\Big)\Big]^{2p} = \frac{m_{2p}(t)}{[m_2(t)]^p}.$$

A candidate $\mathbf{s} \in \mathcal{S}_A$ will be a stationary point of D_{2p} if and only if the directional derivative of $D_{2p}\big(\mathbf{q}(t)\big)$ at $\mathbf{q}(0) = \mathbf{s}$ vanishes in all directions in $\mathcal{S}_A$:

$$\left.\frac{dF_{2p}\big(\mathbf{q}(t)\big)}{dt}\right|_{t=0} = 0, \quad \text{for all } \mathbf{r} \in \mathcal{S}_A.$$

Since $m_2(0) = 1$, a direct calculation gives

$$\begin{aligned}\left.\frac{dF_{2p}}{dt}\right|_{t=0} &= m'_{2p}(0) - p\, m_{2p}(0)\, m'_2(0) \\ &= 2p\Big(\sum_k s_k^{2p-1}(r_k - \langle \mathbf{s}, \mathbf{r}\rangle s_k)\Big) \\ &= 2p\left\langle \mathbf{s}^{\odot(2p-1)}, \mathbf{r} - \langle \mathbf{s}, \mathbf{r}\rangle \mathbf{s}\right\rangle\end{aligned}$$

Equating this to zero results in the orthogonality condition $\mathbf{s}^{\odot(2p-1)} \perp (\mathbf{r} - \langle \mathbf{s}, \mathbf{r}\rangle \mathbf{s})$.

Let now $\mathcal{S}_A^\perp$ denote the orthogonal complement to the set of attainable responses, and likewise decompose $\mathcal{S}_A$ into a one-dimensional subspace colinear with a candidate $\mathbf{s} \in \mathcal{S}_A$ and its resulting orthogonal complement:

$$\begin{aligned}\mathcal{S}_{A\mathbf{s}} &\triangleq \{\mathbf{x} \in \mathcal{S}_A : \mathbf{x} = \alpha \mathbf{s} \text{ for some scalar } \alpha\} \\ \mathcal{S}_{A\mathbf{s}}^\perp &\triangleq \{\mathbf{x} \in \mathcal{S}_A : \mathbf{x} \perp \mathbf{s}\}.\end{aligned}$$

The entire vector space then admits the orthogonal decomposition $\mathcal{S}_{A\mathbf{s}} \oplus \mathcal{S}_{A\mathbf{s}}^\perp \oplus \mathcal{S}_A^\perp$.

Now, for all $\mathbf{r} \in \mathcal{S}_A$, the term $\mathbf{r} - \langle \mathbf{s}, \mathbf{r}\rangle \mathbf{s}$ is restricted to $\mathcal{S}_{A\mathbf{s}}^\perp$, because $\mathbf{s}$ has unit ℓ_2 norm. The orthogonality condition $\mathbf{s}^{\odot(2p-1)} \perp (\mathbf{r} - \langle \mathbf{s}, \mathbf{r}\rangle \mathbf{s})$ then holds if and only if the projection of $\mathbf{s}^{\odot(2p-1)}$ along $\mathcal{S}_{A\mathbf{s}}^\perp$ vanishes. This amounts to saying that the projection of $\mathbf{s}^{\odot(2p-1)}$ onto $\mathcal{S}_A$ reduces to its projection onto $\mathcal{S}_{A\mathbf{s}}$:

$$\mathcal{P}_A\left(\mathbf{s}^{\odot(2p-1)}\right) = \alpha \mathbf{s}, \quad \text{for some scalar } \alpha.$$

With $\mathbf{s}$ scaled to unit norm, finally, we see that

$$\begin{aligned}\alpha &= \alpha\langle \mathbf{s}, \mathbf{s}\rangle = \langle \mathbf{s}, \mathcal{P}_A(\mathbf{s}^{\odot(2p-1)})\rangle \\ &= \langle \mathbf{s}, \mathbf{s}^{\odot(2p-1)}\rangle \qquad \text{because } \mathbf{s} \in \mathcal{S}_A \\ &= \sum_k s_k^{2p} = [D_{2p}(\mathbf{s})]^{2p},\end{aligned}$$

which completes the proof. $\diamond$

For the sufficient order case, $\mathcal{P}_A = \mathbf{I}$, and the statement of Theorem 1 reduces to $\mathbf{s}^{\odot(2p-1)} - \alpha \mathbf{s} = \mathbf{0}$, or

$$s_k(s_k^{2p-2} - \alpha) = 0, \qquad \text{for all } k.$$

This says that all nonzero terms have the same amplitude ($= \sqrt[2p-2]{\alpha}$). For the undermodeled case ($\mathcal{P}_A \neq \mathbf{I}$), we instead have $\mathbf{s}^{\odot(2p-1)} - \alpha \mathbf{s} = \mathbf{b}$ for some $\mathbf{b} \in \mathcal{S}_A^\perp$. Since $\mathbf{b}$ is not known a priori, this relation does not reveal the form of $\mathbf{s}$. Accordingly, the next section develops an iterative procedure which converges to a local maximum of $D_{2p}(\mathbf{s})$ within $\mathcal{S}_A$.

We may remark that the maxima of the limiting criterion $D_\infty(\mathbf{s})$, obtained by letting $p \to \infty$, admit a simple analytic characterization: Each maximum of D_∞ yields (to within an arbitrary scale factor) a Wiener combined response of the form $\mathbf{s} = \mathcal{P}_A(\mathbf{e}_n)$. A compact derivation of this result is submitted in [8].

4 CONSTRUCTION OF MAXIMA OF $D_{2p}(\mathbf{s})$

The stationary points fulfilling Theorem 1 may be identified as the fixed points of a nonlinear map $\mathbf{q} = T_{2p}(\mathbf{s})$, which sends a unit sphere of $\mathcal{S}_A$ to itself, defined as follows:

1. Take $\mathbf{s} \in \mathcal{S}_A$, scaled to unit norm: $\|\mathbf{s}\| = 1$ (the choice of norm is arbitrary for now);
2. Project its Hadamard exponential onto $\mathcal{S}_A$:
$$\mathbf{v} = \mathcal{P}_A(\mathbf{s}^{\odot(2p-1)});$$
3. Scale the result to unit norm: $\mathbf{q} = \mathbf{v}/\|\mathbf{v}\|$ (using the same norm as in step 1).

It is straightforward to check that the fixed points of this map (i.e., those unit-norm $\mathbf{s}$ in $\mathcal{S}_A$ for which $\mathbf{q} = T_{2p}(\mathbf{s}) = \mathbf{s}$) are precisely the stationary points fulfilling Theorem 1. The following inequality applies whenever $\mathbf{s}$ is not a fixed point.

Theorem 2 *Let $\mathbf{s} \in \mathcal{S}_A$ be scaled to unit norm. If $\mathbf{q} = T_{2p}(\mathbf{s}) \neq \mathbf{s}$, then $D_{2p}(\mathbf{q}) > D_{2p}(\mathbf{s})$.*

Before presenting a proof, two remarks are in order:

Remark 1: It follows readily that the iterative procedure

$$\mathbf{s}_{(i+1)} = T_{2p}(\mathbf{s}_{(i)}), \tag{1}$$

in which the subscript (i) denotes the iteration number, will approach a local maximum of $D_{2p}(\mathbf{s})$, save for an exceptional set of initial conditions on $\mathbf{s}_{(0)}$.[1] Which maximum is approached depends on the initialization.

Remark 2: In the sufficient order case ($\mathcal{P}_A = \mathbf{I}$), the iteration (1) reduces to the super-exponential algorithm proposed by Shalvi and Weinstein [5]; if the term of largest amplitude of $\mathbf{s}_{(0)}$ is in position n, then successive iterates $\mathbf{s}_{(i)}$ converge to $\mathbf{e}_n$ at a super-exponential rate. (If one considers the ℓ_∞ norm in steps 1 and 3, then $\|\mathbf{e}_n - \mathbf{s}_{(i+1)}\|_\infty = \|\mathbf{e}_n - \mathbf{s}_{(i)}\|_\infty^{2p-1}$, which is contractive because if $\mathbf{s}_0$ has its term of largest amplitude in position n, and is scaled such that $\|\mathbf{s}_{(0)}\|_\infty = 1$, then $\|\mathbf{e}_n - \mathbf{s}_{(0)}\|_\infty < 1$). For the undermodeled case ($\mathcal{P}_A \neq \mathbf{I}$), the iteration (1) may be found in [5, eqs. (22)–(24)], proposed therein as an approximation within $\mathcal{S}_A$ to the super-exponential convergence algorithm; see also [6]. Shalvi and Weinstein argue that if $\mathcal{P}_A(\mathbf{s}_{(i)}^{\odot(2p-1)}) \approx \mathbf{s}_{(i)}^{\odot(2p-1)}$ for all i, then the trajectory in question should still converge to something resembling a Wiener response. Not specified in [5], though, is how the approximation $\mathcal{P}_A(\mathbf{s}_{(i)}^{\odot(2p-1)}) \approx \mathbf{s}_{(i)}^{\odot(2p-1)}$ should be quantified to ensure convergence. The proof of Theorem 2 to follow, and hence the convergence inference of the previous remark, appeals to no approximation.

[1] The exceptional set will include, e.g., all saddle points and all crest lines leading to such saddle points.

Proof: For convenience, we assume ℓ_2 normalization in the algorithm: $\|\mathbf{s}\|_2 = \|\mathbf{q}\|_2 = 1$. The inequality $D_{2p}(\mathbf{q}) > D_{2p}(\mathbf{s})$ will then follow by showing that

$$\sum_k |q_k|^{2p} > \sum_k |s_k|^{2p}, \quad \text{whenever } \mathbf{q} = T_{2p}(\mathbf{s}) \neq \mathbf{s}.$$

We begin with the identity

$$\begin{aligned} &\sum_k |q_k|^{2p} - \sum_k |s_k|^{2p} \\ &\quad = \langle \mathbf{q}^{\odot(2p-1)}, \mathbf{q}\rangle - \langle \mathbf{s}^{\odot(2p-1)}, \mathbf{s}\rangle \\ &\quad = \langle \mathbf{q}^{\odot(2p-1)} - \mathbf{s}^{\odot(2p-1)}, \mathbf{q}\rangle + \langle \mathbf{s}^{\odot(2p-1)}, \mathbf{q}-\mathbf{s}\rangle \quad (2) \end{aligned}$$

The proof proceeds in two steps:

1. We shall first show the generic inequality
$$\langle \mathbf{q}^{\odot(2p-1)} - \mathbf{s}^{\odot(2p-1)}, \mathbf{q}\rangle \geq (2p-1)\langle \mathbf{s}^{\odot(2p-1)}, \mathbf{q}-\mathbf{s}\rangle$$
valid for any two vectors $\mathbf{q}$ and $\mathbf{s}$. This will then give, with respect to (2),
$$\sum_k |q_k|^{2p} - \sum_k |s_k|^{2p} \geq 2p\,\langle \mathbf{s}^{\odot(2p-1)}, \mathbf{q}-\mathbf{s}\rangle$$

2. We shall then recognize that
$$\langle \mathbf{s}^{\odot(2p-1)}, \mathbf{q}-\mathbf{s}\rangle > 0$$
whenever $\mathbf{q} = T_{2p}(\mathbf{s}) \neq \mathbf{s}$, by virtue of $\mathbf{q}$ being a scaled projection of $\mathbf{s}^{\odot(2p-1)}$.

For the first part, let $\mathbf{x} = \mathbf{q} - \mathbf{s}$; a direct calculation shows that

$$\begin{aligned} &\langle \mathbf{q}^{\odot(2p-1)} - \mathbf{s}^{\odot(2p-1)}, \mathbf{q}\rangle - (2p-1)\langle \mathbf{s}^{\odot(2p-1)}, \mathbf{q}-\mathbf{s}\rangle \\ &\quad = \sum_k \Big((s_k + x_k)^{2p} - s_k^{2p} - 2p|s_k|^{2p-1} x_k \Big) \\ &\quad = \sum_k F(s_k, x_k) \qquad (3) \end{aligned}$$

in terms of the two-variable polynomial

$$F(s,x) = (s+x)^{2p} - s^{2p} - 2p\, s^{2p-1} x.$$

We show now that $F(s,x) \geq 0$ for all real s and x. Let $f(s,x)$ denote the partial derivative with respect to x:

$$\begin{aligned} f(s,x) &\triangleq \frac{\partial F(s,x)}{\partial x} \\ &= 2p\left((s+x)^{2p-1} - s^{2p-1}\right) \begin{cases} > 0, & x > 0; \\ < 0, & x < 0. \end{cases} \end{aligned}$$

We observe that $F(s,0) = 0$ for all s, so that

$$F(s,x) = \begin{cases} \displaystyle\int_0^x f(s,\xi)\,d\xi, & x > 0; \\ \displaystyle -\int_x^0 f(s,\xi)\,d\xi, & x < 0. \end{cases}$$

Since the sign of the integrand is the sign of x, we obtain $F(s,x) \geq 0$ for all s and x. The sum (3) is thus comprised of nonnegative terms, which gives the first part of the proof.

For the second part, let $\mathbf{r}$ vary over all vectors in $\mathcal{S}_A$ of unit ℓ_2 norm. It is straightforward to check that the maximum of $\langle \mathbf{s}^{\odot(2p-1)}, \mathbf{r}\rangle$ is attained if and only if $\mathbf{r} = \mathbf{v}/\|\mathbf{v}\|_2$, where $\mathbf{v} = \mathcal{P}_A(\mathbf{s}^{\odot(2p-1)})$. Since $\mathbf{q}$ is precisely the scaled projection $\mathbf{v}/\|\mathbf{v}\|_2$, while $\mathbf{s} \neq \mathbf{q}$ is not, the second part now follows, to complete the proof. $\diamond$

5 CONCLUDING REMARKS

We have derived a characterization of stationary points for a family of blind equalization criteria in undermodeled cases. The characterization is set in the combined response space, whixh is a more relevant thanthe equalizer coefficient space. The family includes the Godard and Shalvi-Weinstein algorithms, and also leads to the first convergence proof for super-exponential algorithms in undermodeled cases. Extensions of these results to noisy, multi-source channels are under development.

References

[1] Y. Li and Z. Ding, "Global convergence of fractionally spaced Godard (CMA) adaptive equalizers," *IEEE Trans. Sig. Proc.*, vol. 44, pp. 818–826, April 1996.

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

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

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

[5] O. Shalvi and E. Weinstein, "Super-exponential methods for blind deconvolution," *IEEE Trans. Information Theory*, vol. 39, pp. 504–519, March 1993.

[6] Z. Ding, "On convergence analysis of fractionally spaced adaptive blind equalizers," *IEEE Trans. Signal Processing*, vol. 45, pp. 650–657, March 1997.

[7] D. Donoho, "On minimum entropy deconvolution," in: *Applied Time Series Analysis II*, D. F. Findley, ed., Academic, New York, 1981.

[8] P. A. Regalia and M. Mboup, "Undermodeled equalization: A characterization of stationary points for a family of blind criteria," submitted to the *IEEE Trans. Signal Processing*, January 1998.