ON THE CONVERGENT POINTS OF BLIND ADAPTIVE IIR EQUALIZERS

Phillip A. Regalia

Institut National des Télécommunications
9, rue Charles Fourier
F-91011 Evry cedex, France
Phillip.Regalia@int-evry.fr

ABSTRACT

Design procedures for decision feedback equalizers often assume that past decisions are corect, leading to suboptimal performance once incorrect past decisions are injected into the feedback loop. Here we revisit the approach advocated by Labat *et al.* [3], which consists in designing an IIR equalizer using a blind equalization criterion. Instead of adapting the poles and zeros by different criteria, as proposed in [3], we consider the Godard/Shalvi-Weinstein criterion for IIR equalizers. A characterization of stationary points is obtained for the general undermodeled channel case, and simulation results are included.

1. INTRODUCTION

Decision feedback equalization is recognized as a powerful technique to combat intersymbol and co-channel interference [1], [2]. Traditionally the design equations for the feedforward and feedback parts of the system assume that past decisions are correct, or equivalently, that a training sequence is available, leading to suboptimal designs once incorrect past decisions are reinjected into the system.

A potentially more robust approach (e.g., [3], [4], [5]) is to directly design an IIR equalizer based on a blind equalization criterion, such as the Godard/Shalvi-Weinstein criterion. This avoids assuming past decisions to be correct, and the feedback section is optimized at least for a soft-decision model; simulations results from [3] indicate favorable behavior upon replacing the soft decision with a hard decision. Little work has been done, however, on characterizing solutions to blind IIR equalizers in realistic undermodeled cases, due apparently to the nonlinearity of the relevant equations once a higher-order criterion is used for the design of the feedback part [5]. If instead a second-order whitening criterion is used to design the feedback part, as proposed in [3], then the stationary points may be characterized as in [4].

Here we characterize stationary points for adaptive IIR equalizers designed using the Godard/Shalvi-Weinstein criterion, when applied to noisy, multi-user channels, with no knowledge of the channel impulse response length. For ease of notation, z (rather than z^{-1}) will denote the unit delay operator.

2. PROBLEM STRUCTURE

We consider a multi-source communication channel, in which the received signal $\{\mathbf{u}_n\}$, comprising P components, assumes the form

$$(P\updownarrow)\quad \mathbf{u}_n = \sum_{k=0}^{\infty} \mathbf{H}_k\,\mathbf{s}_{n-k} \tag{1}$$

Here the vector $\mathbf{s}_n$ collects Q source signals, which are assumed i.i.d., mutually independent, and scaled to unit variance, while $\{\mathbf{H}_k\}$ is the impulse response of the Q-input/P-output channel. We suppose (realistically) that the number of sources Q exceeds the number of sensors (or oversampling factor) P, giving thus a rectangular channel which will not admit a left inverse.

A seeming generalization to the above model is to include channel noise $\mathbf{b}_n$, leading to the form

$$\mathbf{u}_n = \sum_{k=0}^{\infty} \mathbf{H}_k\,\mathbf{s}_{n-k} + \mathbf{b}_n. \tag{2}$$

If the noise $\mathbf{b}_n$ is Gaussian, however, then we may always develop an innovations model of the form

$$\mathbf{b}_n = \sum_{k=0}^{\infty} \mathbf{F}_k \alpha_{n-k} \tag{3}$$

in which the components of α_n are white, Gaussian, mutually uncorrelated (and hence, i.i.d.), and scaled to unit variance, while $\{\mathbf{F}_k\}$ is the (matrix-valued) impulse response of a modelling filter for the Gaussian process $\mathbf{b}_n$. This, in effect, expresses the noise term as the output of a virtual channel with impulse response $\{\mathbf{F}_k\}$, driven by virtual sources contained in α_n. We may therefore combine

(3) with (2) to write

$$\mathbf{u}_n = \sum_{k=0}^{\infty} \begin{bmatrix} \mathbf{H}_k & \mathbf{F}_k \end{bmatrix} \begin{bmatrix} \mathbf{s}_{n-k} \\ \alpha_{n-k} \end{bmatrix}$$

This assumes the same form as our original model (1) modulo some obvious notational adjustments. (This will generally lead to $Q > P$, i.e., a channel with more inputs than outputs). We may therefore treat background noise and interfering sources without particular distinction, and return to our original model (1).

The matrix $\mathbf{D}$ will denote the diagonal matrix containing the fourth-order cumulants of the sources; Gaussian noise innovation sources will, of course, have fourth-order cumulants equal to zero. The diagonal operator

$$\mathcal{D} = \begin{bmatrix} \mathbf{D} & & \bigcirc \\ & \mathbf{D} & \\ \bigcirc & & \ddots \end{bmatrix}, \qquad [\infty \times \infty],$$

concatenates copies of $\mathbf{D}$ along the main diagonal.

For the equalizer, we shall consider a P-input/single-output IIR filter of the form sketched in Figure 1, *viz.*

$$\begin{aligned} \mathbf{g}(z) &= \sum_{k=0}^{\infty} \mathbf{g}_k z^k = \frac{\begin{bmatrix} B_1(z) & B_2(z) & \cdots & B_P(z) \end{bmatrix}}{A(z)} \\ &= \frac{\mathbf{b}_0 + \mathbf{b}_1 z + \cdots + \mathbf{b}_N z^N}{1 + a_1 z + \cdots + a_M z^M} \end{aligned}$$

expressed using a common denominator $A(z) = 1 + a_1 z + \cdots + a_M z^M$ which is assumed minimum phase (all zeros in $|z| > 1$), and where each numerator polynomial $B_i(z)$ has degree note exceeding N. The P-input/single-output equalizer impulse response is denoted $\{\mathbf{g}_k\}$.

The equalizer output y_n may now be written in many equivalent ways as

$$\begin{aligned} y_n &= \mathbf{g}(z)\,\mathbf{u}_n = \sum_{k=0}^{\infty} \mathbf{g}_k\, \mathbf{u}_{n-k} \quad \text{(equalizer)} \\ &= \mathbf{b}(z) \underbrace{\frac{\mathbf{u}_n}{A(z)}}_{\triangleq\, \mathbf{x}_n} = \sum_{k=0}^{N} \mathbf{b}_k\, \mathbf{x}_{n-k} \quad \text{(partial regressor)} \\ &= \mathbf{g}(z)\,\mathbf{H}(z)\,\mathbf{s}_n = \sum_{k=0}^{\infty} \mathbf{c}_k\, \mathbf{s}_{n-k} \quad \text{(combined response)} \end{aligned}$$

The final line uses the *combined response* $\{\mathbf{c}_k\}$ which is obtained as the convolution of the channel and equalizer impulse responses. The previous line uses the *partially filtered regressor* $\mathbf{x}_n$ obtained by passing the received signal $\mathbf{u}_n$ through the all-pole filter $1/A(z)$ (cf. Figure 1). We may write this partially filtered regressor in terms of the source sequence as

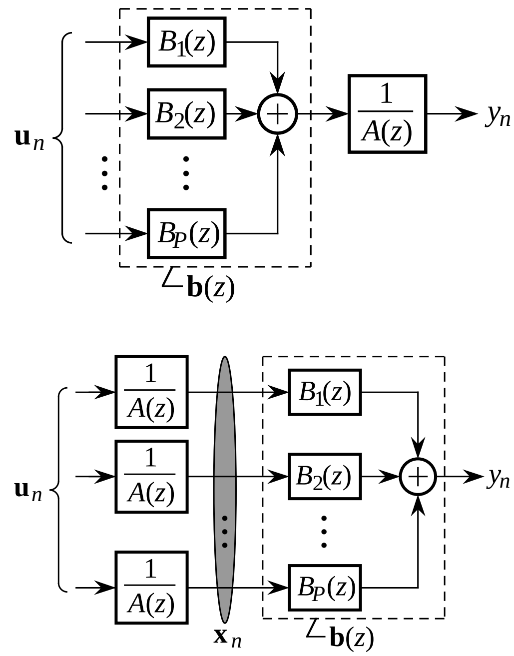

Figure 1: IIR equalizer, using two cascade arrangements.

$$\mathbf{x}_n = \sum_{k=0}^{\infty} \Phi_k\, \mathbf{s}_{n-k}, \quad \text{where} \quad \sum_{k=0}^{\infty} \Phi_k\, z^k = \frac{\mathbf{H}(z)}{A(z)}$$

3. MAIN RESULTS

We shall characterize the stationary points of the Shalvi-Weinstein criterion

$$J[A(z), \mathbf{b}(z)] = \frac{\operatorname{cum}_4(y_n)}{[E(y_n^2)]^2} \tag{4}$$

where $\operatorname{cum}_4(y_n) = \operatorname{cum}(y_n, y_n, y_n, y_n)$ is the fourth-order cumulant of the equalizer output. Note that each stationary point (resp., minimum) of this function in the parameter space generates a stationary point (resp., minimum) of the constant modulus criterion in the same parameter space, and vice-versa [3], [6], [7]. Therefore, our characterization of stationary points summarized below applies to the constant modulus algorithm as well.

To this end, let

$$\mathbf{c} = \begin{bmatrix} \mathbf{c}_0 & \mathbf{c}_1 & \mathbf{c}_2 & \cdots \end{bmatrix}^t$$

be the vector containing the combined response sequence.

Similarly, let

$$\mathcal{F} = \begin{bmatrix} \Phi_0^t & \bigcirc & \cdots & \bigcirc \\ \Phi_1^t & \Phi_0^t & \ddots & \vdots \\ \Phi_2^t & \Phi_1^t & \Phi_0^t & \bigcirc \\ \vdots & \ddots & \ddots & \vdots \end{bmatrix} \quad [\infty \times P(N+1)]$$

be the convolution matrix associated with the partial channel $\Phi(z) = \mathbf{H}(z)/A(z)$. In this way, we may write

$$\mathbf{c} = \mathcal{F} \begin{bmatrix} \mathbf{b}_0^t \\ \vdots \\ \mathbf{b}_N^t \end{bmatrix}$$

in terms of the numerator coefficients $\{\mathbf{b}_k\}$. Note that the combined response always lies in the column space of $\mathcal{F}$, but that this column space varies with $A(z)$. We shall let $\mathcal{P}_A$ denote the orthogonal projector onto this column space:

$$\mathcal{P}_A = \mathcal{F}(\mathcal{F}^t\mathcal{F})^{-1}\mathcal{F}^t$$

Finally, we shall need the autocorrelation and cumulant slice sequences:

$$\begin{aligned} r_k &= E(y_n y_{n-k}) \\ \gamma_k &= \text{cum}(y_n, y_n, y_n, y_{n-k}) \end{aligned} \qquad k = 0, 1, 2, \ldots,$$

as well as their one-sided transforms

$$\begin{aligned} R(z) &= \frac{r_0}{2} + \sum_{k=1}^{\infty} r_k z^k, \qquad |z| < 1 \\ \Gamma(z) &= \frac{\gamma_0}{2} + \sum_{k=1}^{\infty} \gamma_k z^k, \qquad |z| < 1. \end{aligned}$$

If we scale the combined response to unit ℓ_2 norm ($\|\mathbf{c}\|_2 = 1$), then $r_0 = E(y_n^2) = 1$ and $\gamma_0 = \text{cum}_4(y_n) = J$ from (4), and we have:

Theorem 1 *Let* $\mathbf{c}$ *be scaled to unit norm.*

1. *The derivatives of J with respect to the coefficients of* $\mathbf{b}(z)$ *all vanish if and only if*

$$\mathcal{P}_A \mathcal{D} \mathbf{c}^{\odot 3} = \gamma_0 \mathbf{c}$$

where $\mathbf{c}^{\odot 3}$ *is the Hadamard (or componentwise) cube of the vector* $\mathbf{c}$*, i.e.,* $[\mathbf{c}^{\odot 3}]_k = c_k^3$.

2. *The derivatives with respect to the denominator coefficients* $\{a_i\}$ *all vanish if and only if*

$$\gamma_0 R(z_i) = \Gamma(z_i), \qquad i = 1, 2, \ldots, M;$$

where the $\{z_i\}_{i=1}^M$ *are the reciprocals of the zeros of* $A(z)$.

Proof: For part 1, if we fix the coefficients $\{a_k\}$ and absorb the factor $1/A(z)$ into the channel, we are left with an adjustable FIR equalizer $\mathbf{b}(z)$. The characterization of stationary points from [9] then applies immediately.

For part 2, the derivatives of J with respect to the coefficients $\{a_k\}$ become, for $k = 1, 2, \ldots, M$,

$$\begin{aligned} \frac{dJ}{da_k} &= \frac{d}{da_k} \left. \frac{\text{cum}(y_n, y_n, y_n, y_n)}{[E(y_n^2)]^2} \right|_{E(y_n^2)=1} \\ &= 4\left[\text{cum}\left(\frac{dy_n}{da_k}, y_n, y_n, y_n \right) - \gamma_0 E\left(\frac{dy_n}{da_k} y_n \right) \right]. \end{aligned}$$

Now, with $y_n = [\mathbf{b}(z)/A(z)]H(z)\mathbf{s}_n$, we have

$$\frac{dy_n}{da_k} = \frac{-z^k}{A(z)} \frac{\mathbf{b}(z)}{A(z)} H(z)\mathbf{s}_n = \frac{-z^k}{A(z)} y_n, \qquad k = 1, 2, \ldots, M.$$

Therefore, with the expansion $1/A(z) = \sum_{i=0}^{\infty} \alpha_i z^i$, the M constraints $dJ/da_k = 0$ may be written as

$$\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & \alpha_0 & \alpha_1 & \cdots & \alpha_{M-1} & \cdots \\ 0 & 0 & \alpha_0 & \cdots & \alpha_{M-2} & \ddots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \ddots \\ 0 & 0 & \cdots & 0 & \alpha_0 & \cdots \end{bmatrix}}_{\mathcal{C}} \left(\begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \vdots \end{bmatrix} - \gamma_0 \begin{bmatrix} r_0 \\ r_1 \\ r_2 \\ \vdots \end{bmatrix} \right)$$

From the Beurling-Lax theorem [10], the null space of $\mathcal{C}$ is composed of all vectors $\mathbf{f} = [f_0, f_1, f_2, \ldots]^t$ whose strictly causal part yields a z-transform $f_+(z) = \sum_{i=1}^{\infty} f_i z^i$ which vanishes at the reciprocals of the zeros of $A(z)$. Setting $f(z) = \Gamma(z) - \gamma_0 R(z)$, we have $f(0) = \gamma_0 - \gamma_0 r_0 = 0$ since we have scaled the combined response such that $r_0 = E(y_n^2) = 1$, giving $f(z)$ strictly causal. Part 2 now follows from the relation $f(z_i) = 0$ for $i = 1, 2, \ldots, M$. $\diamond$

Note that $R(z)$ is a positive real function ($\text{Re}\,R(e^{j\omega}) \geq 0$ for all ω) because $\{r_k\}$ is an autocorrelation sequence. The scaled function $\Gamma(z)/\gamma_0$, however, will not be positive real in general. Part 2 above shows, interestingly, that stationary points with respect to the denominator coefficients $\{a_i\}$ result in $\Gamma(z)/\gamma_0$ interpolating the positive real function $R(z)$ at the reciprocals of the poles of the equalizer. This then implies that the Pick matrix $\mathbf{P}$ with (i, j)-element

$$\mathbf{P}_{ij} = \frac{\Gamma(z_i) + \Gamma^*(z_j)}{\gamma_0(1 - z_i z_j^*)}$$

is positive definite (e.g., [11], [12]).

4. ADAPTATION ALGORITHMS

Expand each numerator polynomial as

$$B_p(z) = b_0^p + b_1^P z + b_2^p z^2 + \cdots + b_N^p z^N, \qquad p = 1, 2, \ldots, P.$$

The general form for the constant modulus algorithm is

$$\begin{aligned} b_k^p(n+1) &= b_k^p(n) - \mu(y_n^2-1)y_n \frac{dy_n}{db_k^p} \\ a_k(n+1) &= a_k(n) - \mu(y_n^2-1)y_n \frac{dy_n}{da_k} \end{aligned}$$

By writing the equalizer output as

$$y_n = \frac{[B_1(z)\ B_2(z)\ \cdots\ B_P(z)]}{A(z)} \underbrace{\begin{bmatrix} u^1(n) \\ u^2(n) \\ \vdots \\ u^P(n) \end{bmatrix}}_{\mathbf{u}_n}$$

the gradient signals are deduced readily as

$$\begin{aligned} \frac{dy_n}{db_k^p} &= \frac{z^k}{A(z)} u^p(n) \\ \frac{dy_n}{da_k} &= \frac{-z^k}{A(z)} y_n \end{aligned}$$

The adaptation algorithm then appears as

$$\begin{aligned} b_k^p(n+1) &= b_k^p(n) - \mu(y_n^2-1)y_n x_k^p(n) \\ a_k(n+1) &= a_k(n) + \mu(y_n^2-1)y_n \xi_k(n) \end{aligned}$$

in which the computation of the filtered regressors $x_k^p(n)$ and $\xi_k(n)$ is illustrated in Figure 2, for the order $M = N = 3$.

A direct-form filter can go unstable once its coefficients are adapted, unless special constraints are absorbed, such as $\sum_k |a_k(n)| \leq \beta < 1$ for all n [13]. This instability problem can be avoided by using a recursive lattice filter instead [10]; a flowgraph realization appears as Figure 3, with the corresponding adaptation equations as

$$\begin{aligned} b_k^p(n+1) &= b_k^p(n) - \mu(y_n^2-1)y_n x_k^p(n) \\ \theta_k(n+1) &= \theta_k(n) + \mu(y_n^2-1)y_n \xi_k(n) \end{aligned}$$

Simulation: Consider a binary source passing through a single-input, two-output channel $\begin{bmatrix} H_1(z) \\ H_2(z) \end{bmatrix}$, whose impulse response, arranged as two vectors, is

$$\mathbf{h}_1 = \begin{bmatrix} -0.158210 \\ -0.092443 \\ -0.456130 \\ -0.427693 \\ -0.264448 \\ -0.034160 \\ -0.012634 \end{bmatrix}, \quad \mathbf{h}_2 = \begin{bmatrix} 0.025340 \\ -0.375847 \\ -0.797623 \\ -0.931303 \\ -0.667453 \\ -0.302098 \\ -0.070648 \end{bmatrix}.$$

For this example, the two channels $H_1(z)$ and $H_2(z)$ share common (minimum phase) zeros at $z = 1 \pm j$, thereby violating the channel disparity condition. White Gaussian

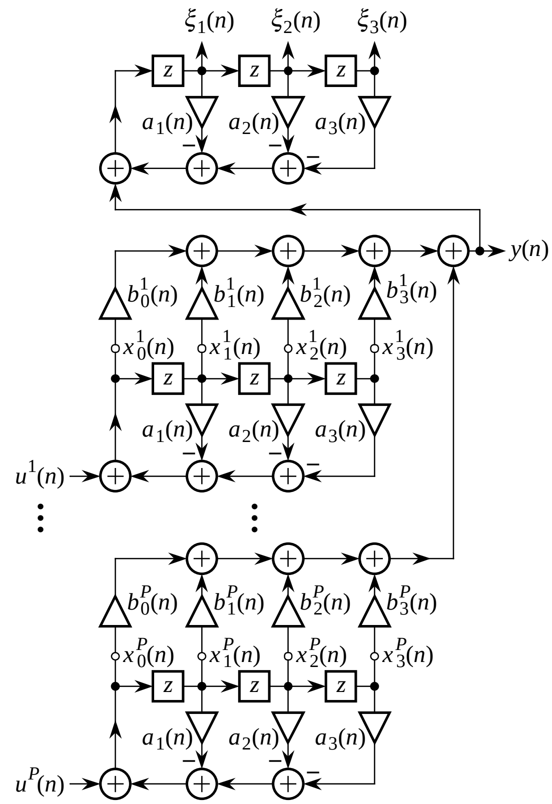

Figure 2: Filtered regressor signals for the IIR constant modulus algorithm in direct form.

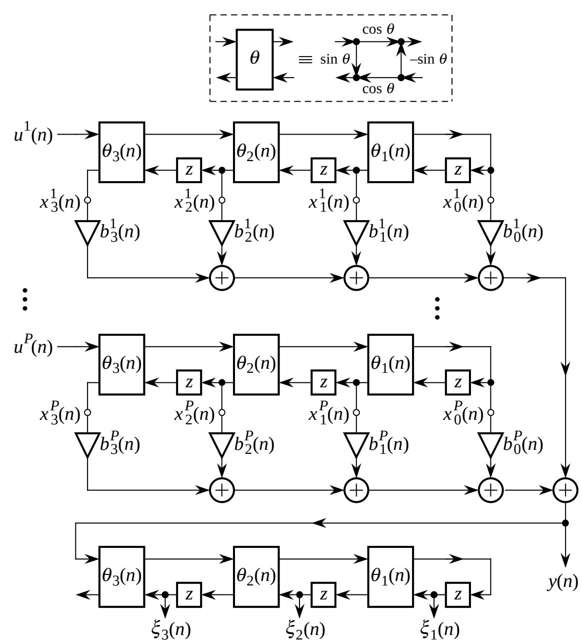

Figure 3: Filtered regressor signals for the IIR constant modulus algorithm in lattice form.

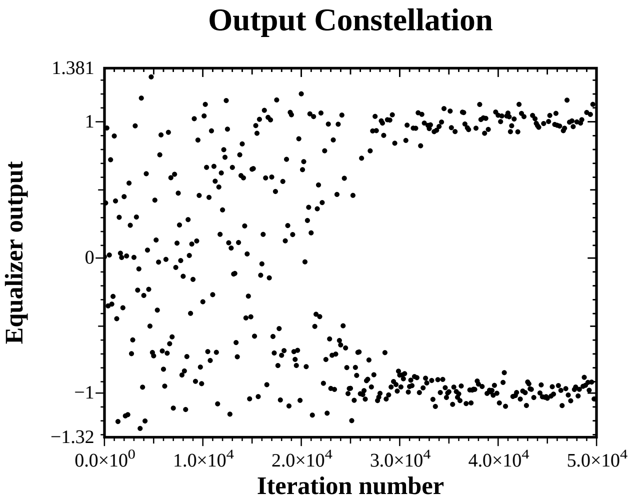

Figure 4: Decimated output sequence from the lattice IIR constant modulus algorithm for the simulation example.

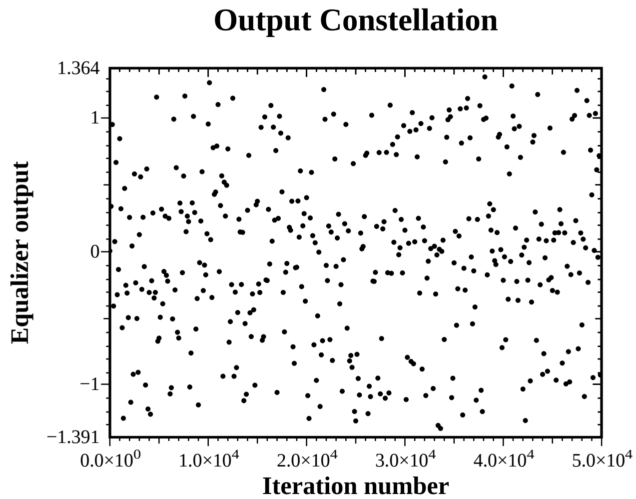

Figure 5: Decimated output sequence from the standard FIR constant modulus algorithm for the same simulation example.

noise, of variance -40 dB, is added to either channel output. Figure 4 shows the (decimated) output sequence generated from the lattice IIR constant modulus algorithm, using $N = M = 4$ for the filter order and $\mu = 0.003$ for the step size, initialized with a sole nonzero coefficient $b_3^1 = 1$. Figure 5 shows the (decimated) output from the standard FIR constant modulus algorithm, using the same parameters and initialization. The inability of the FIR version to open the eye is not surprising, given that the channel disparity condition is violated.

For channels which satisfy cleanly the disparity condition (no common zeros), on the other hand, the FIR version of the Godard algorithm performs quite well [6].

5. REFERENCES

[1] J. G. Proakis, *Digital Communications*, 3rd ed., McGraw-Hill, New York, 1995.

[2] C. A. Belfiore and J. H. Parks, Jr., "Decision feedback equalization," *Proc. IEEE*, vol. 67, August 1979.

[3] J. Labat, O. Macchi, and C. Laot, "Adaptive decision feedback equalization: Can you skip the training period?" *IEEE Trans. Communications*, vol. 46, pp. 921–930, July 1998.

[4] R. López-Valcarce and R. Pérez-González, "On blind adaptive algorithms for IIR equalizers," *DSP 2000*, Hunt, TX, October 2000.

[5] A. Touzny, L. Tong, R. Casas, and C. R. Johnson, Jr., "On AR equalization with the constant modulus criterion," *Proc. EUSIPCO-2000*, Tampere, Finland, September 2000.

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

[7] P. A. Regalia, "On the equivalence between the Godard and Shalvi-Weinstein schemes of blind equalization," *Signal Processing*, vol. 73, pp. 185–190, February 1999.

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

[9] P. A. Regalia and M. Mboup, "Equalization in noisy multi-user channels," *Proc. IEEE/EURASIP Workshop Nonlinear Signal and Image Processing*, Antalya, Turkey, June 1999.

[10] P. A. Regalia, *Adaptive IIR Filtering in Signal Processing and Control*, Marcel Dekker, New York, 1995.

[11] Ph. Delsarte, Y. Genin, and Y. Kamp, "On the role of the Nevanlinna-Pick problem in circuit and system theory," *Int. J. Circuit Theory and its Applications*, vol. 9, pp. 177–187, 1981.

[12] H. Dym, *J-Contractive Matrix Functions, Reproducing Kernel Hilbert Spaces, and Interpolation*, American Mathematical Society, Providence, RI, 1989.

[13] P. Bauer, M. Mansour, and J. Duran, "Stability of polynomials with time-variant coefficients," *IEEE Trans. Circuits and Systems I*, vol. 40, pp. 423–426, June 1993.