ON BLIND ADAPTIVE ALGORITHMS FOR IIR EQUALIZERS

Roberto López-Valcarce

Dept. Electrical and Computer Engineering
University of Iowa, IA, USA
valcarce@icaen.uiowa.edu

Fernando Pérez-González

Dept. Tecnoloxías das Comunicacións
Universidade de Vigo, Vigo, Spain
fperez@tsc.uvigo.es

ABSTRACT

Previous work has shown that the optimum MSE causal equalizer is IIR, with a number of poles equal to the channel order. We investigate two different blind criteria for the adaptation of the recursive part of the equalizer: Output Variance Minimization (OVM) and a Pseudolinear Regression (PLR) method. In sufficient order cases both algorithms converge to the desired setting. In undermodeled cases (i.e. the number of poles in the equalizer is less than the channel order) these algorithms do not necessarily converge to a MSE minimum, but they generally provide acceptable performance. It is shown that under mild conditions PLR always admits a stationary point.

1. INTRODUCTION

Adaptive equalizers are usually implemented using FIR filters, since these systems are unconditionally stable.However, it is known [7] that the optimum Mean-Squared Error (MSE) linear equalizer hass in general an IIR transfer function. In many cases the performance improvement obtained by using an IIR equalizer with respect to that of an FIR structure with the same number of coefficients can be significant.

In [7] the adaptation of the IIR equalizer was done by means of the Kalman filtering algorithm. This requires knowledge about the channel, which was obtained via an identification stage using a training signal. A different alternative was suggested in [3], which presented an unsupervised adaptation mechanism exploiting the structure of the equalization problem. In particular, the adaptations of the recursive and nonrecursive parts of the equalizer can be decoupled and carried out under different criteria.

Our goal is to provide an analysis of the criterion used for the recursive part. Two different scenarios appear: the 'sufficient order' and the 'reduced order' cases. In the former the order of the recursive filter matches the length of the channel. Two possible adaptation rules are then possible: output variance minimization (OVM) by means of a gradient descent, and a pseudolinear regression (PLR) algorithm, both of which converge to the optimum filter. In the reduced order case the situation is more complex. We provide some initial results about the general behavior of these algorithms, concerning issues such as characterization, existence, and uniqueness of stationary points. It is shown via examples that, although OVM and PLR need not provide optimum equalizers in general, the loss in performance is usually small for moderate degrees of undermodeling. This loss is the price to pay for the use of unsupervised criteria.

2. PROBLEM SETTING

Figure 1 depicts the equalization problem. The transmitted symbols $\{w_n\}$ and the noise $\{\eta_n\}$ are assumed stationary, independent, and white with variances σ_w^2 and σ_η^2. The channel has transfer function $H(z)$. The received signal is $u_n = H(z)w_n + \eta_n$, which has autocorrelation $r_u[k] = E[u_n u_{n-k}]$ and psd $S_u(z) = \sum_{-\infty}^{\infty} r_u[k]z^{-k}$.

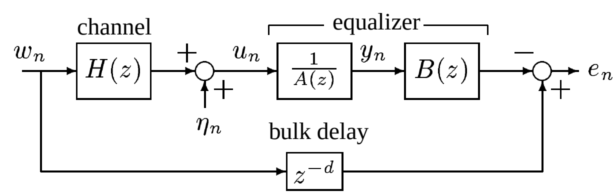

Figure 1: Channel-equalizer configuration

The equalizer transfer function is $B(z)/A(z)$. It was shown in [7] that for FIR channels $H(z) = h_0 + h_1 z^{-1} + \cdots + h_M z^{-M}$ the equalizer that minimizes the MSE $E[e_n^2]$ with e_n as in figure 1 has M poles and d zeros, where d is the associated delay:

$$\frac{B(z)}{A(z)} = \frac{b_0 + b_1 z^{-1} + \cdots + b_d z^{-d}}{1 + a_1 z^{-1} + \cdots + a_M z^{-M}}.$$

In [3], the feedforward part $B(z)$ is adapted blindly by means of the Constant Modulus Algorithm (CMA). We focus on the adaptation of the recursive part $1/A(z)$.

3. THE TWO BLIND ALGORITHMS

Let $\bar{\mathbf{y}}_n = [\ y_n \quad y_{n-1} \quad \cdots \quad y_{n-d}\ ]^T$ and $\mathbf{b} = [\ b_0 \quad b_1 \quad \cdots \quad b_d\ ]^T$. The error e_n can be written as $e_n = w_{n-d} - \mathbf{b}^T\bar{\mathbf{y}}_n$, so that the optimal value of $\mathbf{b}$ is $\mathbf{b}_* = \mathbf{R}_y^{-1}\mathbf{p}_y$, where

$$\mathbf{R}_y = E[\bar{\mathbf{y}}_n\bar{\mathbf{y}}_n^T], \qquad \mathbf{p}_y = E[w_{n-d}\bar{\mathbf{y}}_n]. \tag{1}$$

Thus assuming that $B(z)$ is optimized as a function of $A(z)$, one obtains a 'reduced error surface' of the form

$$J = E[e_n^2]_{\mathbf{b}=\mathbf{b}_*} = \sigma_w^2 - \mathbf{p}_y^T\mathbf{R}_y^{-1}\mathbf{p}_y. \tag{2}$$

Observe that J is a function of $A(z)$ alone. As shown in [7], if $H(z)$ is FIR with degree M, then J is minimized when $y_n = \frac{1}{A(z)}u_n$ is white. That is, the optimum $A(z)$ is the minimum phase spectral factor of $S_u(z)$. This observation provides a blind criterion (whitening $\{y_n\}$) for the adaptation of $1/A(z)$. Note that direct minimization of J does not lend itself to blind adaptation; however, two unsupervised schemes are available for whitening $\{y_n\}$. The first is the minimization of $E[y_n^2]$ via a gradient descent; we refer to this as the Output Variance Minimization (OVM) criterion. Defining $s_n = \frac{1}{A(z)}y_n$ and $\mathbf{a} = [\ a_1 \quad \cdots \quad a_M\ ]^T$, $\mathbf{s}_{n-1} = [\ s_{n-1} \quad \cdots \quad s_{n-M}\ ]^T$, the OVM algorithm is

$$\mathbf{a}_{n+1} = \mathbf{a}_n + \mu y_n \mathbf{s}_{n-1}, \tag{3}$$

where $\mu > 0$ is a small stepsize. Approximating $s_n \approx y_n$, the pseudolinear regression (PLR) algorithm is obtained:

$$\mathbf{a}_{n+1} = \mathbf{a}_n + \mu y_n \mathbf{y}_{n-1}, \tag{4}$$

where $\mathbf{y}_{n-1} = [\ y_{n-1} \quad \cdots \quad y_{n-M}\ ]^T$. PLR disposes of the additional filter $1/A(z)$ that OVM requires for the computation of s_n. Also, PLR presents improved stability behavior due to its 'self-stabilization' property [5, ch. 15], so that stability monitoring is usually not necessary. This is not the case for OVM, which may easily become unstable when the optimum filter has poles close to the unit circle.

4. STATIONARY POINTS

We proceed now to analyze the stationary points of OVM and PLR. Ideally, these stationary points should provide minimization of J. In general they can be characterized as follows. Let $V(z) = z^{-M}A(z^{-1})/A(z)$ be the all-pass function associated with $A(z)$, and for any function $f(z) = \sum_{k=-\infty}^{\infty} f_k z^{-k}$, let $[f(z)]_+ = f_1 z^{-1} + f_2 z^{-2} + \cdots$ denote be the operator extracting the strictly causal part. Use of the Beurling-Lax theorem [8, ch. 3] shows that $1/A(z)$ is a stationary point of OVM or PLR iff it satisfies, for some causal function $g(z) = \sum_{k=0}^{\infty} g_k z^{-k}$ with $\sum |g_k|^2 < \infty$,

$$\left[\frac{S_u(z)}{A(z)A(z^{-1})}\right]_+ = z^{-1}V(z)g(z) \qquad \text{(OVM)} \tag{5}$$

$$\left[\frac{S_u(z)}{A(z)}\right]_+ = z^{-1}V(z)g(z) \qquad \text{(PLR)} \tag{6}$$

Observe that $\frac{S_u(z)}{A(z)A(z^{-1})}$ is the psd of the process $\{y_n\}$. (6) is the z-domain statement of the conditions

$$E[y_n y_{n-k}] = 0, \qquad 1 \le k \le M, \tag{7}$$

that must be satisfied at any stationary point of PLR. These can be seen as an approximation to the whiteness conditions $E[y_n y_{n-k}] = 0$ for all $k > 0$.

From (5) and (6) we immediately see that the stationary points of the two algorithms do not necessarily coincide. Recall that our principal concern is the minimization of J as given in (2); let us consider the sufficient order case first.

5. THE SUFFICIENT ORDER CASE

If $H(z)$ is FIR with the same degree M as $A(z)$ (or less), then $\{u_n\}$ is a moving average process of order M, or MA(M). Then as stated above, J is minimized if $\{y_n\}$ is white. The following is a consequence of [1] and [4]:

Lemma 1 *If $\{u_n\}$ is an MA process of order M or less, then both (5) and (6) have a unique minimum phase solution $A(z)$ which is the minimum phase spectral factor of $S_u(z)$.*

Thus in sufficient order settings both OVM and PLR present a single stationary point which coincides with the minimum of J. This is an additional advantage of the blind criteria: observe that the cost J is a highly nonlinear function of $A(z)$ which could present local minima.

In addition, this stationary point is locally convergent for both algorithms. This is clear for OVM since the optimum is a minimum of $E[y_n^2]$. For PLR, the argument is more subtle; see [9].

6. REDUCED ORDER CASE: OVM

When $\{u_n\}$ is not MA(M), no $A(z)$ with degree M exists that completely whitens $\{y_n\}$. In that case there is no simple description for the minimizer of J. Nevertheless, the cost $E[y_n^2]$ can still be seen as a proxy to J, as we now discuss. Consider the case $d = 0$. Then $\mathbf{R}_y = E[y_n^2]$, $\mathbf{p}_y = E[w_n y_n]$ are both scalars. The reduced error surface becomes

$$J = \sigma_w^2 - \frac{(E[w_n y_n])^2}{E[y_n^2]}, \qquad (d = 0).$$

Note that, since $\{w_n\}$ is white, the term $E[w_n y_n]$ does not depend on $A(z)$ at all, so that minimization of J is equivalent to the minimization of $E[y_n^2]$, i.e. the OVM criterion.

When $d > 0$ this is not strictly true. However, note that

$$J = \sigma_w^2 - \mathbf{p}_y^T \mathbf{R}_y^{-1} \mathbf{p}_y \leq \sigma_w^2 - \frac{||\mathbf{p}_y||^2}{\lambda_{\max}}, \quad (8)$$

where $\lambda_{\max}$ is the largest eigenvalue of $\mathbf{R}_y$. Thus in order to minimize J it makes sense to make $\lambda_{\max}$ as small as possible; and since $\lambda_{\max} \geq \mathrm{trace}(\mathbf{R}_y)/(d+1) = E[y_n^2]$, by making $E[y_n^2]$ small one could expect to decrease $\lambda_{\max}$. This argument is loose since $\mathbf{p}_y$ depends on $A(z)$ for $d > 0$. The examples in section 8 will illustrate this point.

7. REDUCED ORDER CASE: PLR

The first question that arises about the behavior of PLR in undermodeled settings is the existence of stationary points. This issue is not trivial since PLR does not correspond to the minimization of any meaningful cost. To this purpose, we present a new approach which reveals the stationary points of PLR as fixed points of an off-line iterative scheme. First observe that for fixed $A(z)$, one has $y_n = u_n - \mathbf{a}^T \mathbf{y}_{n-1}$. Therefore the conditions (7) can be rewritten as

$$E[\mathbf{y}_{n-1}\mathbf{y}_{n-1}^T]\mathbf{a} = E[u_n \mathbf{y}_{n-1}]. \quad (9)$$

Note that this equation is not linear in $\mathbf{a}$ since $\mathbf{y}_{n-1}$ depends on $\mathbf{a}$, though it suggests the following iterative process:

1. At iteration i, let $\mathbf{a} = \mathbf{a}^{(i)}$ be fixed and let $\mathbf{y}_{n-1}^{(i)} = [\ y_{n-1}^{(i)} \ \cdots \ y_{n-M}^{(i)}\ ]^T$ with $y_n^{(i)} = u_n - \sum_{j=1}^{M} a_j^{(i)} y_{n-j}^{(i)}$.
2. Let $\mathbf{a}^{(i+1)} = E\left[\mathbf{y}_{n-1}^{(i)}(\mathbf{y}_{n-1}^{(i)})^T\right]^{-1} E\left[u_n \mathbf{y}_{n-1}^{(i)}\right]$.
3. Iterate Steps 1 and 2 until convergence.

Clearly, at any fixed point of this iteration, (9) is satisfied, so that these fixed points coincide with the stationary points of PLR. The following result holds now:

Theorem 1 *If the psd $S_u(z)$ is bounded and nonzero for all $|z| = 1$, then the iterative method admits a fixed point $\mathbf{a}_*$ corresponding to a minimum phase polynomial $A_*(z)$.*

Observe that any small amount of white measurement noise in $\{u_n\}$ will yield $S_u(z) \neq 0$ for all $|z| = 1$. Theorem (1) then shows that under these mild conditions the PLR adaptive algorithm admits at least one stationary point inside the stability region. In some cases this point is unique:

Lemma 2 *The minimum phase stationary point of PLR is unique if either (1) $M = 1$, i.e. $A(z) = 1 + a_1 z^{-1}$, or (2) $S_u(z) = \frac{c^2}{(1-\alpha z^{-p})(1-\alpha z^p)}$ with $p = 1$ or $p = 2$.*

Whether this stationary point is *always* unique remains an open question. For example, if $\{u_n\}$ is MA($M+1$), it can be shown that the conditions (6) become

$$\left[\frac{S_u(z)}{A(z)}\right]_+ = r_u[M+1]z^{-1}V(z). \quad (10)$$

Although (10) is highly structured, showing the uniqueness of its solution (if the case) for $M > 1$ seems difficult.

Observe that the above iterative off-line scheme resembles in a sense the Steiglitz-McBride (SM) system identification method, in that a nonlinear problem is replaced by an iteration of linear ones. Some results have been derived in order to lower bound the performance of SM stationary points [8]. We are currently investigating the possibility of using that approach in order to come up with a priori upper bounds on the cost J (or $E[y_n^2]$) at stationary points of PLR.

8. EXAMPLES

Here we illustrate the main points with a few reduced order examples. In all of them we take $\sigma_w^2 = 1$ and $\sigma_\eta^2 = 0.1$.

Example 1: Let $M = 1$ and $H(z) = \frac{1}{1+pz^{-1}}$, $|p| < 1$. With $d = 1$, J is unimodal for $|p| < \frac{\sqrt{5}-1}{2}$ and bimodal for $\frac{\sqrt{5}-1}{2} < |p| < 1$. OVM presents a single minimum, which is the solution of $a_1^2 p(1 + a_1 p) = a_1 + p$ with $|a_1| < 1$, while the unique stationary point of PLR is $a_1 = -p$. Figure 2 shows the variation of the global minimum of J and the OVM and PLR solutions as a function of p, together with the normalized loss $\frac{J-J_{\min}}{J_{\min}}$. For $p = 0$ we recover the sufficient order case. The degree of undermodeling increases with $|p|$, with a corresponding degradation in performance.

Example 2: Let $M = 1$ again, but now $H(z) = 1 + z^{-1} + qz^{-2}$. The reduced cost J for $d = 1$ is unimodal for all values of q. Figure 3 shows the variation of its global minimum, the OVM and PLR solutions, and the normalized excess error, for $|q| \leq 2$. In this example the OVM and PLR solutions remain close to the global minimum of J for $q > 0$, while for $q < 0$ the performance loss is larger.

Example 3: Let $M = 2$ and $H(z) = 1 + z^{-1} + 1.5z^{-2} + 0.5z^{-3} + 0.2z^{-4}$. Figure 4 shows the contour plots of J for $d = 1, 2, 3$, together with those of OVM (which in this case is unimodal) and the PLR stationary point, in the stability domain $|k_i| < 1$ ($k_1 = \frac{a_1}{1+a_2}$, $k_2 = a_2$). For $d > 1$ J becomes multimodal. Even though OVM and PLR do not exactly minimize J, they still provide good performance.

9. CONCLUSIONS

An analysis of two blind criteria (OVM and PLR) for adaptive IIR equalizers has been presented, with emphasis in the undermodeled case. Although in general these algorithms do not converge to an MSE minimum, they usually provide

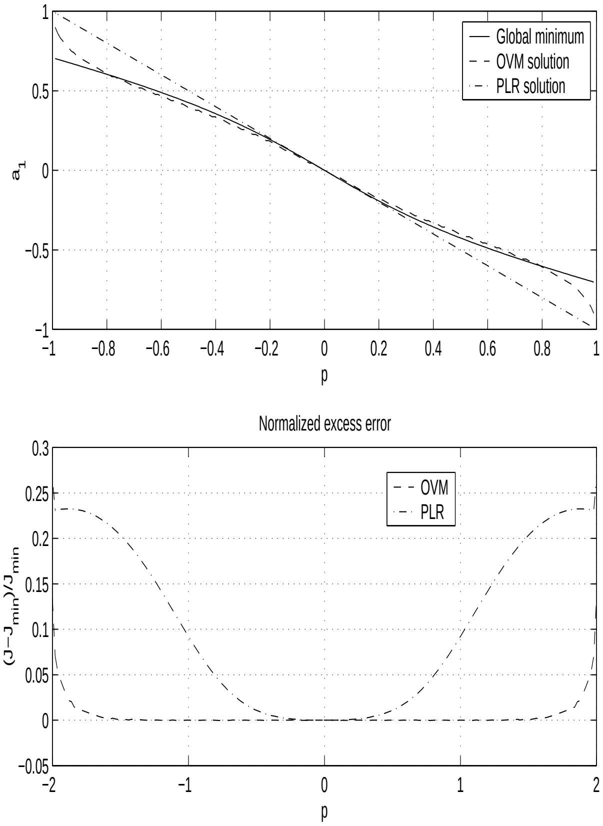

Figure 2: Performance of OVM and PLR for Example 1.

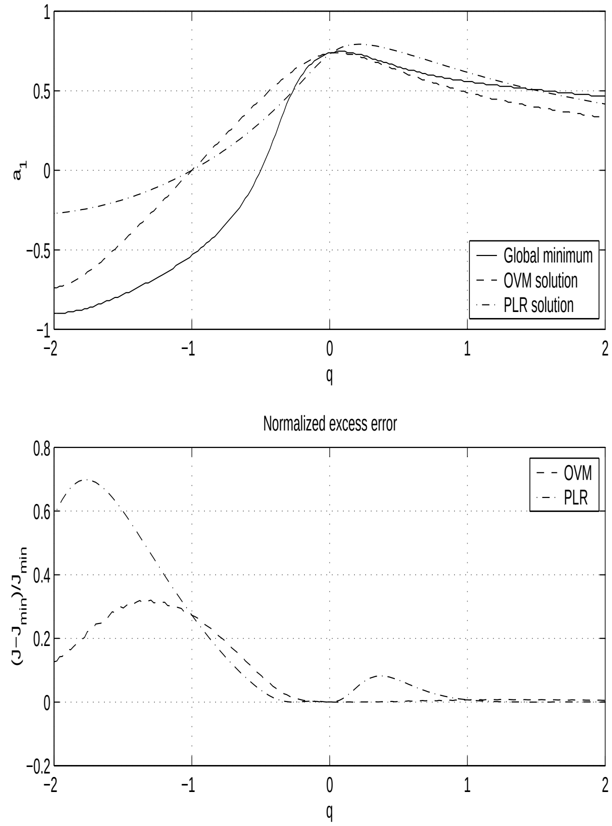

Figure 3: Performance of OVM and PLR for Example 2.

acceptable performance. A characterization of the stationary points has been given, and it was shown that PLR always admits a stationary point, which in some cases is unique. The questions of uniqueness in the general case, lower bounds on performance, and local convergence, remain open.

10. APPENDIX

Proof of Theorem 1: Define the function $F : R^M \to R^M$ as $F(\mathbf{a}) = E[\mathbf{y}_{n-1}\mathbf{y}_{n-1}^T]^{-1}E[u_n\mathbf{y}_{n-1}]$. Then the iterative method can be written as $\mathbf{a}^{(i+1)} = F(\mathbf{a}^{(i)})$. Introduce now the map $\mathbf{k} = L(\mathbf{a})$, where $\mathbf{k} = [\ k_1 \ \ \cdots \ \ k_M\ ]^T$ is the vector of reflection coefficients (lattice parameters) associated to the polynomial $A(z)$, and which can be obtained via the Schur recursion [8]. The stability domain becomes the following convex open subset of R^M:

$$\mathcal{D} = \{\ \mathbf{k} \ \ : \ \ |k_i| < 1 \ \text{ for } i = 1, \ldots, M \ \}.$$

We can define the function $G : \mathcal{D} \to R^M$ as the composition $G = L \circ F \circ L^{-1}$. Thus the iteration can be reparameterized in lattice coordinates as $\mathbf{k}^{(i+1)} = G(\mathbf{k}^{(i)})$. Observe that the iteration may break if $\mathbf{k}^{(i)}$ is outside $\mathcal{D}$.

Let $\partial\mathcal{D}$ denote the boundary of $\mathcal{D}$. As the vector of lattice parameters $\mathbf{k}$ approaches $\partial\mathcal{D}$, at least one root of the corresponding $A(z)$ approaches the unit circle. In that case the diagonal entries of the matrix $E[\mathbf{y}_{n-1}\mathbf{y}_{n-1}^T]$ become

$$E[y_n^2] = E\left[\left(\frac{1}{A(z)}u_n\right)^2\right] \to \infty, \qquad (11)$$

because $S_u(z) > 0$ on the unit circle is assumed. On the other hand the components of the vector $E[u_n\mathbf{y}_{n-1}]$ are given by

$$E[u_n y_{n-j}] = \frac{1}{2\pi i}\oint_{|z|=1} S_u(z)\frac{z^{j-1}}{A(z^{-1})}dz \qquad (12)$$

for $1 \le j \le M$. For minimum phase $A(z)$, the only poles of the integrand inside the unit circle are those of $S_u(z)$; therefore, using the residue theorem to evaluate (12), we see that these quantities remain finite even as one or more roots of $A(z)$ approach $|z| = 1$.

This implies that as $\mathbf{k} \to \partial\mathcal{D}$ from inside $\mathcal{D}$,

$$F(L^{-1}(\mathbf{k})) \to \mathbf{0},$$

in view of (11), (12) and the definition of F. Since $L(\mathbf{0}) = \mathbf{0}$, we conclude that the domain of G can be extended in order

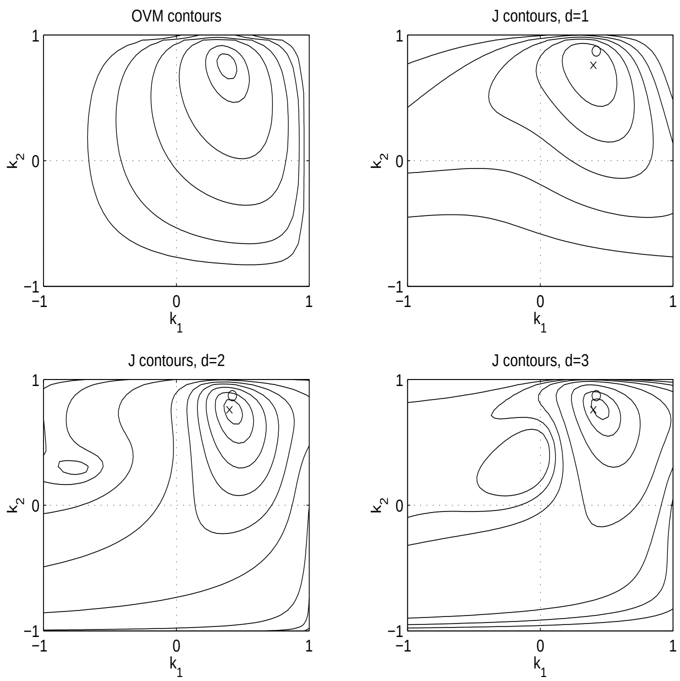

Figure 4: Contours of OVM and J for Example 3. $\circ$ = PLR stationary point; $\times$ = OVM minimum.

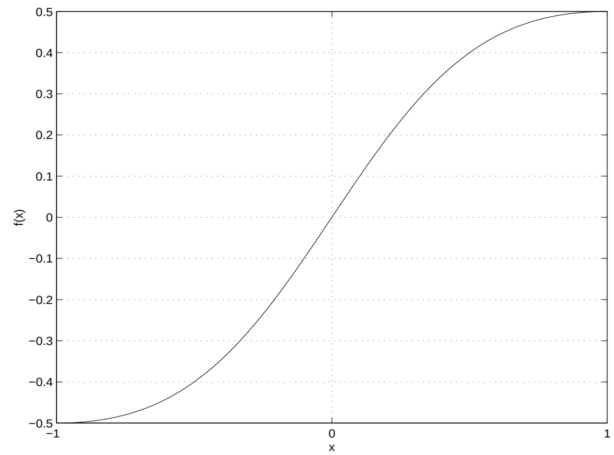

Figure 5: The function $f(x) = x/(1+x^2)$ is one-to-one on $[-1, 1]$.

to include $\partial\mathcal{D}$ by defining $G(\mathbf{k}) = \mathbf{0}$ for all $\mathbf{k} \in \partial\mathcal{D}$, and in this way $G : \mathcal{D} \cup \partial\mathcal{D} \to R^M$ remains continuous.

We now invoke the following Borsuk fixed point theorem [2, p.46]:

Theorem 2 *Let $\bar{\mathcal{D}}$ be a closed, bounded, symmetric and convex subset of R^M, and let G be a continuous mapping from $\bar{\mathcal{D}}$ to R^M. If G is odd along the boundary, i.e.*

$$G(-\mathbf{k}) = -G(\mathbf{k}) \quad \text{for all } \mathbf{k} \in \partial\mathcal{D}, \tag{13}$$

then G admits a fixed point in $\bar{\mathcal{D}}$: there exists $\mathbf{k}_* \in \bar{\mathcal{D}}$ such that $G(\mathbf{k}_*) = \mathbf{k}_*$.

We can apply this result with $\bar{\mathcal{D}} = \mathcal{D} \cup \partial\mathcal{D}$: since $G(\mathbf{k}) = \mathbf{0}$ for all $\mathbf{k} \in \partial\mathcal{D}$, (13) is clearly satisfied. Thus $G(\mathbf{k}_*) = \mathbf{k}_*$ for some $\mathbf{k}_* \in \bar{\mathcal{D}}$. Moreover $\mathbf{k}_*$ cannot lie on $\partial\mathcal{D}$ since all $\partial\mathcal{D}$ is mapped onto $\mathbf{0}$ which is not in $\partial\mathcal{D}$. Thus the fixed point lies inside the stability domain $\mathcal{D}$. ■

Proof of lemma 2(1): With $M = 1$, assume that the polynomials $1 + az^{-1}$ and $1 + bz^{-1}$ were both minimum-phase stationary points of PLR. Define the processes

$$y_n^{(a)} = \frac{1}{1+az^{-1}} u_n, \qquad y_n^{(b)} = \frac{1}{1+bz^{-1}} u_n,$$

and their correlation coefficients

$$r_a[k] = E[y_n^{(a)} y_{n-k}^{(a)}], \qquad r_b[k] = E[y_n^{(b)} y_{n-k}^{(b)}].$$

Then we must have $r_a[1] = r_b[1] = 0$. Define the process $\{x_n\}$ as

$$x_n = \frac{1}{1+az^{-1}} y_n^{(b)} = \frac{1}{1+bz^{-1}} y_n^{(a)},$$

with correlation coefficients $r_x[k] = E[x_n x_{n-k}]$. Noting that $y_n^{(a)} = (1 + bz^{-1})x_n$, it can be easily shown that

$$r_a[1] = r_x[1] + (r_x[0] + r_x[2])b + r_x[1]b^2. \tag{14}$$

Define the cross-correlation coefficients $c[k] = E[y_n^{(b)} x_{n-k}]$. Since $x_n = y_n^{(b)} - ax_{n-1}$, they satisfy

$$c[k] = r_b[k] - ac[k+1], \tag{15}$$
$$r_x[k] = c[k] - ar_x[k-1]. \tag{16}$$

By (15), $r_b[1] = 0$ implies $c[1] = -ac[2]$. In view of (16), this gives

$$r_x[1] = -a(r_x[0] + c[2]), \tag{17}$$
$$\begin{aligned} r_x[2] &= c[2] - ar_x[1] \\ &= c[2] + a^2(r_x[0] + c[2]). \end{aligned} \tag{18}$$

Substituting this into (14), we obtain

$$\begin{aligned} r_a[1] &= -(a - (1+a^2)b + ab^2)(r_x[0] + c[2]) \\ &= 0 \quad \text{by assumption.} \end{aligned} \tag{19}$$

Now note that since

$$r_x[0] + r_x[2] = (1 + a^2)(r_x[0] + c[2])$$

we cannot have $r_x[0] + c[2] = 0$ due to the fact that $|r_x[0]| > |r_x[2]|$. Thus (19) reads as $a - (1 + a^2)b + ab^2 = 0$, or equivalently

$$\frac{a}{1+a^2} = \frac{b}{1+b^2}.$$

Observe that the function $f : [-1, 1] \to [-\frac{1}{2}, \frac{1}{2}]$ given by $f(x) = \frac{x}{1+x^2}$ is one-to-one, see Figure 5. Thus $f(a) = f(b)$ implies $a = b$, since $|a| < 1$, $|b| < 1$. ■

Proof of lemma 2(2): If

$$S_u(z) = \frac{c^2}{(1-\alpha z^{-p})(1-\alpha z^p)},$$

then

$$y_n = \frac{1}{A(z)} u_n$$

is an autoregressive (AR) process of order $M + p$. Without loss of generality we can assume $|\alpha| < 1$. Because $\{y_n\}$ is AR($M + p$), the Forward Prediction Error Filter (FPEF) of order $M + p$ associated to the process $\{y_n\}$ has transfer function $(1 - \alpha z^p)A(z)$.

Consider the lattice parameterization of this FPEF, given by the reflection coefficients $k_1, \ldots, k_{M+p}$. If $A(z)$ is a PLR stationary point, then the first M autocorrelation coefficients of $\{y_n\}$ are zero, according to (7). This in turn implies that

$$k_1 = k_2 = \cdots = k_M = 0. \quad (20)$$

(Incidentally, (20) is true regardless of whether $\{y_n\}$ is AR or not: it holds as long as (7) is true, and can be easily shown via the Levinson recursion for the linear prediction problem, e.g. [6]). Now if $p = 1$, (20) means that one must have

$$(1 - \alpha z^p)A(z) = 1 + k_{M+1} z^{-M-1},$$

which reads as

$$\begin{aligned} a_1 - \alpha &= 0, \\ a_i - a_{i-1}\alpha &= 0, \quad 2 \le i \le M, \\ -a_M \alpha &= k_{M+1}. \end{aligned}$$

The solution to these equations is unique and it is given by $A(z) = 1 + \alpha z^{-1} + \cdots + \alpha^M z^{-M}$.

On the other hand, if $p = 2$, then (20) implies

$$(1 - \alpha z^p)A(z) = \\ 1 + (k_{M+1}k_{M+2})z^{-1} + k_{M+1}z^{-M-1} + k_{M+2}z^{-M-2}.$$

Equating the coefficients of equal powers of z,

$$\begin{aligned} a_1 &= k_{M+1}k_{M+2}, \\ a_2 - \alpha &= 0, \\ a_i - a_{i-1}\alpha &= 0, \quad 3 \le i \le M, \\ -a_{M-1}\alpha &= k_{M+1}, \\ -a_M \alpha &= k_{M21}. \end{aligned}$$

Again these equations have a unique solution, which is given by

$$a_i = \begin{cases} 0 & \text{for } i \text{ odd}, \\ \alpha^{i/2} & \text{for } i \text{ even}. \end{cases}$$

■

11. REFERENCES

[1] K. J. Åström and T. Söderström, "Uniqueness of the Maximum Likelihood estimates of the parameters of an ARMA model", *IEEE Trans. Automatic Control*, vol. 19(6), pp. 769–773, Dec. 1974.

[2] J. Dugundi and A. Granas, *Fixed point theory*, Warsaw, Poland: PWN-Polish Scientific, 1982.

[3] J. Labat, O. Macchi and C. Laot, "Adaptive decision feedback equalization: can you skip the training period?", *IEEE Trans. Communications*, vol. 46(7), pp. 921-930, July 1998.

[4] L. Ljung, T. Söderström and I. Gustavsson, "Counterexamples to general convergence of a commonly used recursive identification method", *IEEE Trans. Automatic Control*, vol. 20(5), pp. 643–652, Oct. 1975.

[5] O. Macchi, *Adaptive processing : the LMS approach with applications in transmission*, Wiley, 1995.

[6] J. Makhoul, "Linear prediction: a tutorial review", *Proc. IEEE*, vol. 63(4), pp. 561-580, April 1975.

[7] B. Mulgrew and C. F. N. Cowan, "An adaptive Kalman equalizer: structure and performance", *IEEE Trans. Acoustics, Speech, Signal Processing*, vol. 35(12), pp. 1727-1735, Dec. 1987.

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

[9] P. Stoica, J. Holst and T. Söderström, "Eigenvalue location of certain matrices arising in convergence analysis problems", *Automatica*, vol. 18(4), pp. 487–489, 1982.