Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Analysis of a first-order adaptive recursive predictor

Roberto López-Valcarce

Abstract— **Adaptive all-pole predictors have recently found renewed interest in the area of digital data transmission due to their ability to perform blind magnitude equalization of the communication channel. The pseudolinear regression (PLR) algorithm constitutes an appealing candidate for the predictor update, since it is computationally simpler than its forerunners. We analyze the behavior of a first-order complex-valued PLR-updated predictor to show that the stationary point is unique even in general undermodelled settings, and that the predictor pole will not escape the unit circle for sufficiently slow adaptation. With no undermodelling, global convergence is also established. Additional properties of PLR solutions in undermodelled scenarios are also given, such as expressions for their prediction gain.**

I. INTRODUCTION

A recursive predictor is a pole-zero filter whose purpose is to provide a model for the spectrum of its input signal. Several algorithms have been suggested for the adaptation of the predictor coefficients, the most prominent being the Recursive Maximum Likelihood (RML) and the Pseudolinear Regression (PLR) methods [5]. RML can be seen as a stochastic gradient descent of the cost function given by the average power of the adaptive filter output. To compute the corresponding gradient term, an additional copy of the recursive portion of the filter is required. On the other hand, PLR directly constructs its update term from the internal signals of the predictor, and therefore its computational complexity is smaller than that of RML. This makes PLR the preferred choice in applications where a predictor with a large number of poles is required. However, due to the lack of an optimization criterion underlying the PLR method, its convergence and stability properties are not well understood. This is especially true for undermodelled settings, in which the predictor does not have enough coefficients to accurately model the input spectrum.

In recent years there has been considerable interest in recursive predictors in the area of digital communications, due to the fact that these devices can be incorporated as the front end of blind adaptive equalizers [3], [4], [7]. The purpose of these predictors is to whiten the received signal as a first step in the equalization process, which can be done online with RML or PLR without the need of a training signal. In most practical cases (i) the transmitted symbols are statistically independent, (ii) the discrete-time equivalent channel presents an impulse response of finite duration, and (iii) the additive noise can be assumed white. Therefore the received signal can be modeled as a moving average (MA) process, so that a purely recursive (that is, all-pole) predictor with order no less than that of the discrete-time equivalent channel will be able to provide perfect whitening when properly tuned.

In view of this, a convergence analysis of all-pole predictors updated with the PLR algorithm would be extremely useful. Although a few results are available in the literature, they usually assume a 'sufficient order' setting, and thus do not apply to the case in which the channel order is larger than that of the all-pole predictor (no perfect whitening is then possible). Under this 'sufficient order' condition, it is known that the PLR algorithm applied to all-pole predictors presents a unique stationary point (yielding perfect whitening) [6] which in addition is locally convergent [9] (as a consequence of the all-pole structure of the predictor, no 'positive realness' condition is required of the stationary point [9]). Analysis of the general undermodelled case is difficult, although conditions for the existence of stationary points have been recently given [7].

We present an analysis of PLR for a single-pole predictor. Uniqueness of the stationary point is proven, even in under-modelled settings. Making use of the ordinary differential equation (ODE) method [2] we show the self-stabilizing property of the algorithm, as well as global stability in the sufficient order case. Having in mind the application of these predictors to digital communications, we focus on the general case of complex-valued signals and filters.

II. ALGORITHM DESCRIPTION

Let $\{u_n\}$ be a complex-valued, wide-sense stationary, zero-mean stochastic process with autocorrelation and power spectral density (psd) respectively given by

$$r_u(k) \triangleq E[u_n u_{n-k}^*], \quad S_u(z) \triangleq \sum_{k=-\infty}^{\infty} r_u(k) z^{-k}. \quad (1)$$

The output signal $\{e_n\}$ (the prediction error) of an all-pole predictor of order M is computed as

$$e_n = u_n - \mathbf{a}_n^H \mathbf{e}_n, \quad (2)$$

where the coefficient and regressor vectors are defined respectively as

$$\mathbf{a}_n \triangleq [\ a_1(n) \ \cdots \ a_M(n)\ ]^T, \quad (3)$$

$$\mathbf{e}_n \triangleq [\ e_{n-1} \ \cdots \ e_{n-M}\ ]^T. \quad (4)$$

The PLR algorithm updates the predictor coefficients as

$$\mathbf{a}_{n+1} = \mathbf{a}_n + \mu \mathbf{e}_n e_n^*, \quad (5)$$

Work supported by the *Ramón y Cajal* program of the Spanish Ministry of Science and Technology.

R. López-Valcarce is with the Department of Signal Theory and Communications, Universidad de Vigo, 36310 Vigo, Spain. E-mail: valcarce@gts.tsc.uvigo.es

with $\mu > 0$ a small stepsize. Note that if a stationary point is reached, then the conditions

$$r_e(k) \triangleq E[e_n^* e_{n-k}] = 0, \qquad 1 \le k \le M, \tag{6}$$

must hold. Thus, PLR attempts to achieve whiteness of the prediction error $\{e_n\}$ by nulling out its autocorrelation coefficients of lags 1 through M.

Let $\mathbf{a}_\star$ be a stationary point of (5). Multiplying both sides of (2) by $\mathbf{e}_n^H$ and taking expectations, we obtain

$$\underbrace{E[e_n \mathbf{e}_n^H]}_{=\mathbf{0}} = E[u_n \mathbf{e}_n^H] - \mathbf{a}_\star^H \underbrace{E[\mathbf{e}_n \mathbf{e}_n^H]}_{=r_e(0)\mathbf{I}} \tag{7}$$

Hence, the coefficient and signal vectors at a stationary point must satisfy

$$\mathbf{a}_\star = \frac{1}{r_e(0)} E[u_n^* \mathbf{e}_n]. \tag{8}$$

Variance of prediction error

Using (8), an expression for the prediction error variance can be obtained. Multiplying (2) by e_n^* and taking expectations,

$$\begin{aligned} r_e(0) &= E[u_n e_n^*] - \mathbf{a}_\star^H E[\mathbf{e}_n e_n^*] = E[u_n e_n^*] \\ &= E[|u_n|^2] - E[u_n \mathbf{e}_n^H]\mathbf{a}_\star \qquad (9) \\ &= r_u(0) - (\mathbf{a}_\star^H \mathbf{a}_\star) r_e(0). \qquad (10) \end{aligned}$$

In (9) and (10) we have used (2) and (8) respectively. Therefore, one has

$$r_e(0) = \frac{r_u(0)}{1 + \mathbf{a}_\star^H \mathbf{a}_\star}, \tag{11}$$

which directly relates the prediction error variance to the input variance and the predictor coefficients. Note from (11) that, at any stationary point of PLR, $r_e(0) \le r_u(0)$ holds; in other words, any recursive predictor obtained by the PLR algorithm cannot be worse than the trivial predictor $\mathbf{a} = \mathbf{0}$, in terms of the prediction error variance achieved.

Improvement in spectral flatness

The *spectral flatness measure* [10] of a process $\{x_n\}$ with psd $S_x(z)$ is defined as the ratio of geometric to arithmetic means of $S_x(e^{j\omega})$:

$$\gamma_x^2 \triangleq \frac{\exp\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln S_x(e^{j\omega}) \mathrm{d}\omega\}}{\frac{1}{2\pi}\int_{-\pi}^{\pi} S_x(e^{j\omega}) \mathrm{d}\omega}, \tag{12}$$

which satisfies $0 \le \gamma_x^2 \le 1$ and equals 1 for a flat spectrum (white $\{x_n\}$).

Assume that $S_u(z)$ satisfies the Paley-Wiener condition so that it admits a spectral factorization

$$S_u(z) = K_u B(z) B^*(1/z^*), \tag{13}$$

where $B(z)$ is a minimum phase, monic, rational transfer function, and $K_u > 0$ is the geometric mean of $S_u(e^{j\omega})$.

Let the transfer function of the all-pole predictor be

$$\frac{1}{A(z)} = \frac{1}{1 + a_1^* z^{-1} + \cdots + a_M^* z^{-M}}, \tag{14}$$

and let $S_e(z) = \sum_{k=0}^{\infty} r_e(k) z^{-k}$ be the prediction error psd. Note that

$$S_e(z) = K_u \frac{B(z)}{A(z)} \frac{B^*(1/z^*)}{A^*(1/z^*)} \tag{15}$$

constitutes a spectral factorization of $S_e(z)$. Since $B(z)/A(z)$ is minimum phase, causal and monic, it follows that the geometric mean of $S_e(e^{j\omega})$ equals K_u as well. Therefore the spectral flatness measures at the input and output of the predictor are respectively

$$\gamma_u^2 = \frac{K_u}{r_u(0)}, \qquad \gamma_e^2 = \frac{K_u}{r_e(0)}. \tag{16}$$

From (16) and (11) we conclude that if $\mathbf{a}_\star$ constitutes a stationary point of PLR, the improvement in spectral flatness is given by

$$\frac{\gamma_e^2}{\gamma_u^2} = 1 + \mathbf{a}_\star^H \mathbf{a}_\star \ge 1. \tag{17}$$

Equality holds in (17) if and only if $\mathbf{a}_\star = \mathbf{0}$, which can be a stationary point of PLR if and only if the input process satisfies $r_u(k) = 0$ for $1 \le k \le M$.

III. On uniqueness of the stationary point

Existence of at least one stationary point of PLR in general undermodelled settings, corresponding to a predictor with stable poles, was established in [7][1] under a positivity condition on $S_u(e^{j\omega})$. Whether this point is in general unique remains an open issue. Here we show uniqueness in the $M = 1$, complex-valued case[2].

Theorem 1: Suppose that $M = 1$ and that $S_u(e^{j\omega}) > 0$ for all ω. Then the PLR update rule (5) has a single stationary point inside the stability region $\mathcal{C} = \{z : |z| < 1\}$.

Proof: Let $\{x_n\}$, $\{y_n\}$ be the outputs of two first-order recursive filters with coefficients a, b, driven by $\{u_n\}$:

$$x_n = u_n - a^* x_{n-1}, \quad y_n = u_n - b^* y_{n-1}, \tag{18}$$

with $|a|, |b| < 1$, and let $r_x(\cdot)$, $r_y(\cdot)$ be their respective autocorrelation functions. Define also the process

$$w_n = x_n - b^* w_{n-1}, \tag{19}$$

with autocorrelation $r_w(\cdot)$. In view of (18), $\{w_n\}$ also satisfies

$$w_n = y_n - a^* w_{n-1}. \tag{20}$$

Using (19)-(20), $r_x(1)$ and $r_y(1)$ can be written as

$$\begin{aligned} r_x(1) &= (1 + |b|^2) r_w(1) + b^* r_w(0) + b r_w(2), \qquad (21) \\ r_y(1) &= (1 + |a|^2) r_w(1) + a^* r_w(0) + a r_w(2). \qquad (22) \end{aligned}$$

Suppose that both predictors in (18) are stationary points of (5). Then one has $r_x(1) = r_y(1) = 0$. Thus, from (22),

$$r_w(1) = -\frac{a^* r_w(0) + a r_w(2)}{1 + |a|^2}. \tag{23}$$

[1]Although in [7] real-valued signals and filters were assumed, the extension of the existence proof to the complex-valued case is immediate.

[2]The proof of uniqueness given in [7] for the real-valued, $M = 1$ case does *not* carry over to the complex-valued case.

Substituting (23) into (21) then yields

$$\frac{1+|b|^2}{1+|a|^2}[a^* r_w(0)+a r_w(2)]-[b^* r_w(0)+b r_w(2)]=0. \quad (24)$$

Let $z_0 \triangleq a(1+|b|^2)-b(1+|a|^2)$. Then, if $z_0 \neq 0$, from (24) we could write $r_w(2)/r_w(1)=-z_0^*/z_0$, implying that $|r_w(2)|=|r_w(0)|$. In that case, it is readily found the determinant of the 3×3 autocorrelation matrix of $\{w_n\}$ is nonpositive, which implies that the psd $S_w(z)=\sum_k r_w(k)z^{-k}$ cannot be positive on the whole unit circle. This is a contradiction, since

$$S_w(e^{j\omega})=\frac{S_u(e^{j\omega})}{|1+a^*e^{-j\omega}|^2|1+b^*e^{-j\omega}|^2}>0 \qquad \forall\omega,$$

because $S_u(e^{j\omega})>0$ and $|a|,|b|<1$ by assumption. Hence $z_0=0$, i.e.,

$$\frac{a}{1+|a|^2}=\frac{b}{1+|b|^2}, \quad (25)$$

from which it follows that either $b=a$ or $b=1/a^*$. Since it is assumed that $|a|,|b|<1$, it must hold that $b=a$. ∎

Observe that Theorem 1 holds regardless of whether the stationary point of PLR achieves perfect whitening of the prediction error.

IV. Self-stabilization

Jaidane and Macchi [1] analyzed the behavior of a PLR-updated pole-zero predictor with narrowband inputs. In such setting the poles of the optimal predictor are located on the unit circle $\partial\mathcal{C}=\{z : |z|=1\}$. Adaptation noise will inevitably push the poles *outside* $\partial\mathcal{C}$, and then one could expect the predictor coefficients to 'blow up'. Surprisingly, [1] revealed a 'self-stabilization' mechanism by which excursions outside $\partial\mathcal{C}$ actually make the PLR algorithm push the offending poles back into the stability region $\mathcal{C}$. In another related result, [5, Lemma 4.2] establishes the boundedness of the prediction error in the case of a vanishing stepsize $\mu_n=\bar{\mu}/n$.

Here we consider a (broadband) stochastic input and a single-pole predictor. Our goal is to show that whenever the predictor coefficient a wanders sufficiently close to $\partial\mathcal{C}$, PLR (with constant stepsize) will tend to decrease its magnitude, therefore effectively avoiding filter instability. To do so we exploit the fact that, under slow adaptation (that is, with asufficiently small but not necessarily vanishing stepsize μ), the mean convergence properties of the adaptive algorithm can be studied by examining those of the associated ODE [2], which in this case is

$$\dot{a}(t)=r_e^*(1)|_{a=a(t)}=E[e_n^* e_{n-1}]|_{a=a(t)}, \quad (26)$$

where the right-hand side is evaluated for a fixed parameter $a=a(t)$ in the recursive predictor that generates the prediction error $\{e_n\}$. Then we have the following.

Theorem 2: Suppose that $M=1$ and that the input process psd is bounded away from zero, that is, there exists some $c_1>0$ such that

$$S_u(e^{j\omega})\geq c_1 \qquad \text{for all } \omega. \quad (27)$$

Consider the ODE (26). Then there exists a real constant $c_2<1$ such that the magnitude $|a(t)|$ of the predictor coefficient is a decreasing function of time whenever $c_2<|a(t)|<1$ is satisfied.

Proof: For a fixed single-pole predictor, the prediction error $\{e_n\}$ can be written as

$$e_n=u_n-a^*e_{n-1}=\sum_{k=0}^{\infty}(-a^*)^k u_{n-k}. \quad (28)$$

Therefore the driving term of the ODE (26) satisfies

$$\begin{aligned} r_e^*(1) &= E[u_n^* e_{n-1}]-aE[|e_{n-1}|^2] \\ &= \sum_{k=0}^{\infty}(-a^*)^k r_u^*(k+1)-a r_e(0). \end{aligned} \quad (29)$$

On the other hand, the prediction error variance satisfies

$$\begin{aligned} r_e(0)=E[e_n e_n^*] &= E[u_n e_n^*]-a^*E[e_{n-1}e_n^*] \\ &= \sum_{k=0}^{\infty}(-a)^k r_u(k)-a^* r_e^*(1). \end{aligned} \quad (30)$$

Substituting (30) into (29), one obtains

$$r_e^*(1)=\sum_{k=0}^{\infty}(-a^*)^k r_u^*(k+1)+\sum_{k=0}^{\infty}(-a)^{k+1}r_u(k)+|a|^2 r_e^*(1).$$

With this, the ODE (26) can be written explicitly in terms of a and the input correlation coefficients:

$$\dot{a}=\frac{1}{1-|a|^2}\left(\sum_{k=0}^{\infty}(-a^*)^k r_u^*(k+1)+\sum_{k=0}^{\infty}(-a)^{k+1}r_u(k)\right). \quad (31)$$

We can write the ODE more compactly if we introduce

$$F_u(z)=\frac{r_u(0)}{2}+\sum_{k=0}^{\infty}r_u(k)z^{-k}, \quad (32)$$

such that the input psd satisfies $S_u(z)=F_u(z)+F_u^*(1/z^*)$. Note that on $\mathcal{C}$, one has $S_u(e^{j\omega})=2\,\mathrm{Re}\,F_u(e^{j\omega})$. It is readily found that

$$\begin{aligned} \sum_{k=0}^{\infty}(-a)^k r_u(k+1) &= \frac{1}{-a}\left[F_u(-1/a)-\frac{r_u(0)}{2}\right], \quad (33) \\ \sum_{k=0}^{\infty}(-a)^{k+1}r_u(k) &= -a\left[F_u(-1/a)+\frac{r_u(0)}{2}\right]. \quad (34) \end{aligned}$$

Substituting (33)-(34) into (31) and rearranging terms,

$$\dot{a}=\frac{1}{a^*}\left[\frac{r_u(0)}{2}-\frac{F_u^*(-1/a)+|a|^2F_u(-1/a)}{1-|a|^2}\right]. \quad (35)$$

Consider now the function

$$W(t) \triangleq \frac{1}{2}|a(t)|^2, \quad (36)$$

whose time derivative is $\dot{W}(t)=\mathrm{Re}[a^*(t)\dot{a}(t)]$. Hence, from (35),

$$\dot{W}=\frac{r_u(0)}{2}-\frac{1+|a|^2}{1-|a|^2}\,\mathrm{Re}\,F_u(-1/a). \quad (37)$$

Since $S_u(z)$ satisfies (27), then $\operatorname{Re} F_u(e^{j\omega}) \geq \frac{c_1}{2} > 0$ for all ω, so that $F_u(z)$ is strictly positive real (SPR). As a consequence, it holds that $\operatorname{Re} F_u(z) \geq c_1 > 0$ for all $|z| \geq 1$ [8]. In particular, $\operatorname{Re} F_u(-1/a) \geq c_1$ as long as $|a| \leq 1$.

Let now

$$c_2 = \sqrt{\frac{r_u(0) - c_1}{r_u(0) + c_1}} > 1. \tag{38}$$

Suppose now that at $t = t_0$, a becomes close enough to $\partial\mathcal{C}$ so that $c_2 < |a(t_0)| < 1$ holds. Then it follows that

$$r_u(0) < \frac{1 + |a(t_0)|^2}{1 - |a(t_0)|^2} c_1, \tag{39}$$

and therefore the time derivative (37) satisfies

$$\dot{W}(t_0) \leq \frac{r_u(0)}{2} - \frac{1 + |a(t_0)|^2}{1 - |a(t_0)|^2} \frac{c_1}{2} < 0. \tag{40}$$

Hence $W(t)$ is a decreasing function of time at $t = t_0$, which proves the theorem. ∎

Hence PLR will push the predictor pole inside $\mathcal{C}$ whenever it gets too close to the stability boundary. In fact, we note from (30) and (33)-(34) that the prediction error variance is given by

$$r_e(0) = \frac{2 \operatorname{Re} F_u(-1/a)}{1 - |a|^2}, \tag{41}$$

so that using (41), the time derivative $\dot{W}(t)$ in (37) can be rewritten as

$$\dot{W}(t) = \frac{1}{2}\left[r_u(0) - (1 + |a|^2) r_e(0)\right]. \tag{42}$$

This shows that $|a(t)|$ increases whenever the prediction error variance satisfies

$$r_e(0) < \frac{r_u(0)}{1 + |a|^2},$$

and decreases when the opposite is true. Recall that, from (11), at the stationary point $a_\star$ equality must hold, that is, $r_e(0)|_{a=a_\star} = r_u(0)/(1 + |a_\star|^2)$.

In the general case in which $M > 1$, we can follow the same argument using $W(t) = \frac{1}{2}\mathbf{a}^H\mathbf{a}$, whose time derivative becomes now

$$\begin{aligned} \dot{W}(t) = \operatorname{Re}\{\mathbf{a}^H\dot{\mathbf{a}}\} &= \operatorname{Re}\{\mathbf{a}^H E[\mathbf{e}_n e_n^*]\} \\ &= \operatorname{Re}\{E[(\mathbf{a}^H\mathbf{e}_n) e_n^*]\} \\ &= \operatorname{Re}\{E[u_n - e_n) e_n^*]\} \\ &= \operatorname{Re}\{E[u_n e_n^*]\} - r_e(0). \end{aligned} \tag{43}$$

With the transfer function of the predictor given by (14), the term $E[u_n e_n^*]$ can be expressed as

$$E[u_n e_n^*] = \frac{1}{2\pi j}\oint_{|z|=1} \frac{S_u(z)}{A^*(1/z^*)}\frac{\mathrm{d}z}{z}. \tag{44}$$

Assume all roots of $A(z)$ are inside $\mathcal{C}$. Then, the only poles of the integrand inside $\mathcal{C}$ are those of $S_u(z)/z$. Hence, using the residue theorem to evaluate (44), it is seen that $E[u_n e_n^*]$ will remain finite even if one or more roots of $A(z)$ approach $\partial\mathcal{C}$. In that situation, however, the prediction error variance $r_e(0)$ will grow unbounded (provided $S_u(e^{j\omega}) > 0$ for all ω). Then from (43), we see that eventually $\dot{W}$ will become negative so that the predictor coefficient vector norm will decrease. Unfortunately, this does not necessarily imply that the predictor poles will move away from $\partial\mathcal{C}$, due to the non-spherical shape of the stability region in the space of the filter coefficients when $M > 1$.

V. Global stability in the sufficient order case

If the input $\{u_n\}$ is a Moving Average process of first order, or MA(1), then its psd reduces to

$$S_u(z) = r_u^*(1) z + r_u(0) + r_u(1) z^{-1} \tag{45}$$

so that

$$F_u(z) = r_u(0)\left(\frac{1}{2} + \rho z^{-1}\right) \qquad \text{with} \quad \rho \triangleq \frac{r_u(1)}{r_u(0)}. \tag{46}$$

Note that a necessary and sufficient condition for $F_u(z)$ to be SPR is that $|\rho| < \frac{1}{2}$; and that a first-order recursive predictor is capable of perfectly whitening the prediction error.

We recall that in sufficient order settings (that is, when the input process is MA(M) with M the all-pole predictor order), it is known that the unique stationary point of PLR is locally convergent [9]. In the particular case $M = 1$, global stability can be established, as seen next.

Theorem 3: Suppose that $M = 1$ and that the input is an MA(1) process satisfying $S_u(e^{j\omega}) > 0$ for all ω. Then the unique stationary point in $\mathcal{C}$ of the ODE (26) associated to the PLR update is globally convergent for all $|a(0)| < 1$.

Proof: Substituting the expression for $F_u(-1/a) = r_u(0)(\frac{1}{2} - \rho a)$ in the ODE (35) yields, after some simplification,

$$\dot{a} = r_u(0)\frac{\rho^* - a + \rho a^2}{1 - |a|^2}. \tag{47}$$

The numerator in (47) has two roots $z_{1,2}$ which are conjugate reciprocal, that is, $z_2 = 1/z_1^*$. For $|\rho| < \frac{1}{2}$, one of these roots is always in $\mathcal{C}$, and is given by

$$a_\star = \frac{1 - \sqrt{1 - 4|\rho|^2}}{2|\rho|^2}\rho^*. \tag{48}$$

Therefore, we can rewrite (47) as

$$\begin{aligned} \dot{a} &= r_u(0)\rho\frac{(a_\star - a)(1/a_\star^* - a)}{1 - |a|^2} \\ &= \frac{r_u(0)\rho}{a_\star^*}\frac{(a_\star - a)(1 - a_\star^* a)}{1 - |a|^2} \\ &= \frac{r_u(0)}{1 + |a_\star|^2}\frac{(a_\star - a)(1 - a_\star^* a)}{1 - |a|^2}, \end{aligned} \tag{49}$$

where the last step in (49) follows from the fact that $a_\star$ satisfies $a_\star = (1 + |a_\star|^2)\rho^*$, as can be checked using (48).

To show that the stationary point $a_\star$ of the nonlinear ODE (49) is globally stable, let us introduce the Lyapunov function

$$V(t) \triangleq \frac{1}{2}|a(t) - a_\star|^2, \tag{50}$$

whose time derivative is given by

$$\dot{V}(t) = \operatorname{Re}\{[a(t) - a_\star]^* \dot{a}(t)\}. \tag{51}$$

Now using (49), it is found that

$$(a-a_\star)^*\dot{a} = -\frac{r_u(0)}{1+|a_\star|^2}\frac{|a-a_\star|^2}{1-|a|^2}(1-a_\star^* a), \tag{52}$$

and therefore

$$\begin{aligned}\dot{V}(t) &= -\frac{r_u(0)|a-a_\star|^2}{(1+|a_\star|^2)(1-|a|^2)}\operatorname{Re}(1-a_\star^* a)\\ &= -\frac{r_u(0)|a-a_\star|^2}{(1+|a_\star|^2)(1-|a|^2)}\\ &\quad\times\left[1-|a||a_\star|\cos(\theta-\theta_\star)\right],\end{aligned} \tag{53}$$

where, in polar coordinates,

$$a = |a|e^{j\theta}, \qquad a_\star = |a_\star|e^{j\theta_\star}. \tag{54}$$

It is seen from (53) that

$$|a(t)|<1, \quad t\geq 0 \qquad \Rightarrow \qquad \dot{V}(t)\leq 0, \quad t\geq 0, \tag{55}$$

and that $\dot{V}=0$ if and only if $a=a_\star$. The condition $|a|<1$ will hold provided that $|a(0)|<1$, thanks to the self-stabilizing property of the ODE given in Theorem 2. Therefore, $a(t)\to a_\star$ as $t\to\infty$ for all initializations $|a(0)|<1$. ∎

VI. ORDER M PREDICTOR WITH MA($M+1$) INPUT

Possibly the simplest undermodelled setting one can think of is that in which the input process is MA with degree equal to the order of the all-pole predictor plus one. We now examine the behavior of PLR stationary points under such condition. Thus, if $\{u_n\}$ is an MA($M+1$) process, then its spectral factor $B(z)$ in (13) is a minimum phase polynomial with degree $M+1$.

Let the M-th order all-pole predictor transfer function be $1/A(z)$ as given by (14), and let

$$V(z) = \frac{z^{-M}A^*(1/z^*)}{A(z)} \tag{56}$$

be the corresponding associated allpass transfer function. It was shown in [7] that if $1/A(z)$ is a stable transfer function corresponding to a stationary point of PLR, then

$$\left[\frac{S_u(z)}{A(z)}\right]_+ = z^{-1}V(z)g(z) \tag{57}$$

must hold for some stable and causal $g(z)$, where $[\cdot]_+$ extracts the strictly causal part of its argument. Then, for the MA($M+1$) input case, we have

$$\left[\frac{B(z)}{A(z)}B^*(1/z^*)\right]_+ = z^{-1}V(z)g(z) \tag{58}$$

for some causal, stable $g(z)$. As shown in [8, prob. 8.4], the left-hand side of (58) is a rational function of degree not exceeding that of $B(z)/A(z)$, and any pole of this function is a pole of $B(z)/A(z)$. Therefore we can write

$$\left[\frac{B(z)}{A(z)}B^*(1/z^*)\right]_+ = z^{-1}\frac{q(z)}{A(z)} \tag{59}$$

where $q(z)$ is a polynomial of degree not exceeding M. Equating (58) and (59),

$$\begin{aligned}z^{-1}\frac{q(z)}{A(z)} &= z^{-1}V(z)g(z)\\ &= z^{-1}g(z)\frac{z^{-M}A^*(1/z^*)}{A(z)}.\end{aligned} \tag{60}$$

Since all the roots of $z^{-M}A^*(1/z^*)$ lie outside the unit circle, none of them can be canceled out by a pole of $g(z)$ (since $g(z)$ is causal and stable). Therefore every root of $z^{-M}A^*(1/z^*)$ must also be a root of $q(z)$, and then $g(z)$ must reduce to a constant: $g(z)=g_0$.

With $S_e(z)$ the prediction error psd, note that $S_u(z)/A(z) = S_e(z)A^*(1/z^*)$ and that $S_e(z)$ satisfies

$$S_e(z) = z^{M+1}P^*(1/z^*) + r_e(0) + z^{-(M+1)}P(z), \tag{61}$$

where $P(z)=\sum_{k=0}^{\infty} r_e(k+M+1)z^{-k}$, which is a causal function. This is because at any PLR stationary point the autocorrelation coefficients with lags 1 through M of $\{e_n\}$ are zero. Therefore

$$\begin{aligned}\frac{S_u(z)}{A(z)} &= S_e(z)A^*(1/z^*)\\ &= z^{M+1}P^*(1/z^*)A^*(1/z^*) + r_e(0)A^*(1/z^*)\\ &\quad + z^{-(M+1)}P(z)A^*(1/z^*),\end{aligned} \tag{62}$$

so that the strictly causal part of (62) reduces to $z^{-(M+1)}P(z)A^*(1/z^*)$. Hence, from (60),

$$z^{-(M+1)}P(z)A^*(1/z^*) = z^{-1}g_0\frac{z^{-M}A^*(1/z^*)}{A(z)},$$

which shows that

$$g_0 = r_e(M+1), \qquad P(z) = \frac{r_e(M+1)}{A(z)}.$$

Therefore the psd $S_e(z)$ must take the form

$$S_e(z) = \frac{r_e^*(M+1)z^{M+1}}{A^*(1/z^*)} + r_e(0) + \frac{r_e(M+1)z^{-(M+1)}}{A(z)}, \tag{63}$$

and consequently, since $B(z)B^*(1/z^*) = S_u(z) = S_e(z)A(z)A^*(1/z^*)$, we can write

$$\begin{aligned}S_u(z) &= r_e^*(M+1)z[z^M A(z)]\\ &\quad + r_e(0)A(z)A^*(1/z^*)\\ &\quad + r_e(M+1)z^{-1}[z^{-M}A^*(1/z^*)].\end{aligned} \tag{64}$$

Equating the coefficient of $z^{-(M+1)}$ in both sides of (64), we see that

$$r_u(M+1) = r_e(M+1), \tag{65}$$

that is, the autocorrelation coefficient of lag $M+1$ of $\{e_n\}$ matches that of $\{u_n\}$. This coefficient can be seen as a measure of the degree of undermodeling: if $r_u(M+1)=0$, then the input is MA(M) and the sufficient order setting is recovered.

The relation (64) shows that $A(z)$ is trying to approximate in some sense $B(z)$ in that $r_e(0)A(z)A^*(1/z^*)$ tries to match $B(z)B^*(1/z^*)$, with the additional 'tails' weighted by

$r_e(M+1)$. This shows some degree of robustness of the PLR solution, since for small $|r_e(M+1)|$, $r_e(0)A(z)A^*(1/z^*)$ will be close to $B(z)B^*(1/z^*)$, so that $\{e_n\}$ will be close to white; see (63).

Example: In the real-valued case with $M = 1$ and an MA(2) input process, we can represent (see fig. 1) the prediction gain (17) obtained by PLR as a function of the parameters $\rho_k \triangleq r_u(k)/r_u(0)$, $k = 1, 2$. The admissible values of these parameters are those such that $1 + 2\rho_1 z^{-1} + 2\rho_2 z^{-2}$ is SPR. The maximum values of the prediction gain (less than 3 dB for a first-order predictor) are achieved for input processes with highpass or lowpass spectral densities, such that one zero of the second-order spectral factor $B(z)$ is real and close to $z = \pm 1$, while the other one is real and of the same sign.

Fig. 2 shows the loss in prediction gain of the first-order PLR solution with respect to an optimal predictor (that is, one that truly minimizes the prediction error variance (41)). This loss is within 1 dB for all possible MA(2) inputs.

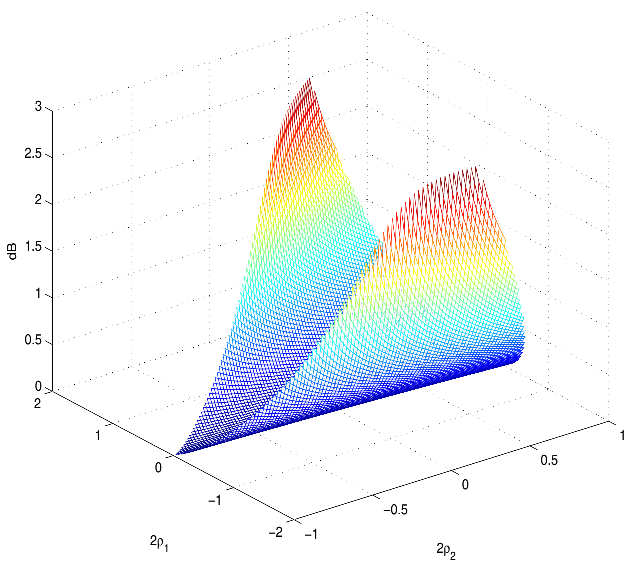

Fig. 1. Prediction gain of a first-order PLR predictor with an MA(2) input process.

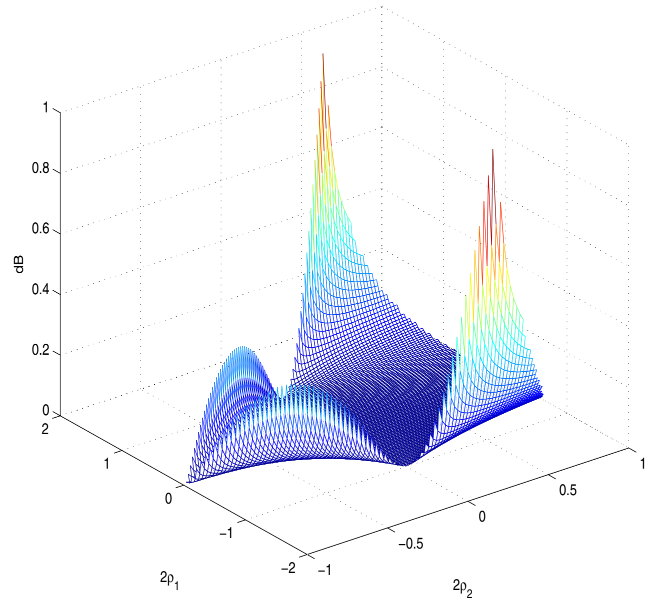

Fig. 2. Loss in prediction gain of a first-order PLR predictor with an MA(2) input process with respect to an optimal first-order predictor.

VII. Conclusions

Due to its low computational cost, the PLR algorithm is an appealing candidate for the adaptation of the all-pole predictor. In the sufficient order case, uniqueness and local convergence of the stationary point are guaranteed. The first-order predictor enjoys a self-stabilizing property, and the solution is unique in the general undermodelled case and globally convergent in the sufficient order case. The prediction loss with respect to a truly optimal predictor seems to be small. Further work should explore the behavior of PLR with higher-order all-pole predictors: for instance, Casas *et al.* [3] have reported a case in which the convergence domain of a third-order predictor in a sufficient-order setting seems to be quite small. Understanding the properties of this adaptive algorithm is crucial if it is to be adopted in a digital communication receiver.

References

[1] M. Jaidane and O. Macchi, 'Quasi-periodic self-stabilization of adaptive ARMA predictors,' *Int. J. Adaptive Control and Signal Processing,* vol. 2, pp. 1-31, Jan. 1988.

[2] H. Fan, 'Application of Benveniste's convergence results in the study of adaptive IIR filtering algorithms,' *IEEE Trans. Information Theory,* vol. 34, pp. 692-709, Jul. 1988.

[3] R. A. Casas, T. J. Endres, A. Touzni, C. R. Johnson, Jr., and J. R. Treichler, 'Current approaches to blind decision feedback equalization,' in *Signal Processing Advances in Communications, Vol. 1: Trends in Channel Estimation and Equalization,* pp. 367-415, Prentice Hall, 2001.

[4] J. Labat, O. Macchi and C. Laot, 'Adaptive decision feedback equalization: can you skip the training period?,' *IEEE Trans. Communications,* vol. 46 no. 7, pp. 921-930, Jul. 1998.

[5] L. Ljung and T. Söderström, *Theory and practice of recursive identification,* MIT Press, Cambridge, 1987.

[6] L. Ljung, T. Söderström, and I. Gustavsson, 'Counterexamples to general convergence of a commonly used recursive identification method,' *IEEE Trans. Automatic Control,* vol. AC-20 no. 5, pp. 643-652, Oct. 1975.

[7] R. López-Valcarce and F. Pérez-González, 'On blind adaptive algorithms for IIR equalizers,' *Proc. 9th Digital Signal Processing Workshop,* Oct. 2000. Online proceedings available at http://spib.ece.rice.edu/SPTM/DSP2000/program.html

[8] P. A. Regalia, *Adaptive IIR filtering in signal processing and control,* Marcel Dekker, New York, 1995.

[9] P. Stoica, J. Holst and T. Söderström, 'Eigenvalue location of certain matrices arising in convergence analysis problems,' *Automatica*, vol. 18 no. 4, pp. 487-489, 1982.

[10] P. P. Vaidyanathan, *Multirate systems and filter banks,* Prentice Hall, Englewood Cliffs, NJ, 1993.