

FILTER DESIGN TO GUARANTEE CONVERGENCE OF THE PSEUDOLINEAR REGRESSION IIR ADAPTIVE ALGORITHM

F. Pérez-González* and C. Abdallah†

* BSP Group, Departamento de Tecnologías de las Comunicaciones, ETSI Telecomunicación, Universidad de Vigo, 36200-VIGO, SPAIN

† ICS Group, Department of Electrical and Computer Engineering University of New Mexico, Albuquerque, NM 87131, USA.

ABSTRACT

In this paper we design a Finite-Impulse-Response (FIR) filter to enforce the so-called Strictly-Positive-Realness (SPR) condition in an adaptive Infinite-Impulse-Response (IIR) filtering context. The paper results may then be used to guarantee the convergence of pseudolinear regression IIR adaptive algorithms.

1. INTRODUCTION

Many of the adaptive IIR filtering approaches can be classified into two categories: (1) the *equation-error formulation* and (2) the *output-error formulation*. Adaptive algorithms based on the output-error formulation are attractive because they produce unbiased solutions when some sufficiency conditions are met. The price is of course, a more complex algorithm due to the nonlinear dependence on the unknown coefficients which implies a multimodal Minimum Square Output Error (MSOE) surface.

Different simplifications have been proposed to mitigate the complexity associated with the *output-error* algorithms. These simplifications usually assume slow parameter adaptation and use the typical gradient-based descent [1]. The pseudolinear regression algorithm (PLR) belongs to this class and corresponds to the following adaptation equation

$$\theta(n+1) = \theta(n) + \alpha R^{-1}(n+1)\phi_0(n)e_0(n) \quad (1)$$

where $\theta(n)$ is the parameter vector, R^{-1} is an estimate of the inverse correlation matrix (usually computed by means of the matrix inversion lemma), $\phi_0(n)$ is the regressor vector whose components are delayed samples of the plant input-output variables and $e_0(n)$ is the output error. Most studies concerning the convergence of the adaptive IIR algorithms start with a system-identification configuration where the desired signal is the output of an unknown IIR filter (plant) whose parameters are to be estimated. It has been shown that for a rich enough input excitation, a sufficient condition for this algorithm to converge, is that the stable denominator polynomial $A(z^{-1})$ of the unknown plant satisfies a strict-positive-realness condition (SPR). The SPR condition stems from hyperstability theory which is a valuable

tool when dealing with time varying systems [2]. The SPR condition is simply formulated for a stable $A(z^{-1})$ as

$$\operatorname{Re}\{A(z^{-1})\} > 0, |z| = 1 \quad (2)$$

Unfortunately, as the degree of $A(z^{-1})$ increases, satisfying condition (2) becomes more critical. If we use $e_f(n)$ instead of $e_o(n)$ in equation (1), where $e_f(n)$ is a filtered version of $e_o(n)$ through an FIR filter $C(z^{-1})$, it can be shown that the SPR condition transforms into

$$\operatorname{Re}\left\{\frac{A(z^{-1})}{C(z^{-1})}\right\} > 0, |z| = 1 \quad (3)$$

where $C(z^{-1})$ is stable. This suggests designing $C(z)$ to guarantee condition (3) and, therefore, global asymptotic convergence of the algorithm [3]. Since $A(z^{-1})$ is usually unknown, one can not choose the naive $C(z) = A(z)$. On the other hand, we can adapt the coefficients of $C(z)$, but the algorithm is then complicated by the requirement that $C(z)$ be minimum-phase, which may require the added complication of projecting the unstable roots of $C(z)$ into the stable region.

The remaining of the paper is organized as follows. Section 2 contains a discussion of the classes of uncertainties often encountered in the adaptive filtering setting, and a review of the importance of the SPR condition. In section 3, we review earlier work that attempted to satisfy the SPR conditions. Section 4 includes a discussion of a simple approach of designing $C(z^{-1})$ to satisfy the SPR condition. The main results are contained in section 5 and consist of an algorithm for the design of a low-order $C(z^{-1})$ which can also satisfy the SPR condition. Our conclusions are given in section 6.

2. UNCERTAINTY AND SPRNESS

In many cases, it is possible to find an *a priori confidence* set where the coefficients of $A(z^{-1})$ are known to lie. This set may have originated from some off-line preliminary identification or, as is sometimes the case, from a physical knowledge of the dynamical behavior of the plant. For instance, the unknown plant may have an impulse response that practically vanishes in a certain amount of time, thus imposing a condition on the placement of the poles. Two approaches

The work of F. Pérez-González was partially supported by Northern Telecom.

have been recently considered to characterize the uncertainty. In the first, uncertainty bounds may be given in the space of the coefficients of $A(z^{-1})$. The resulting set usually has a polytopic form (i.e., the coefficients depend affinely on a number of independent parameters). The main advantage of considering this characterization is the result of Anderson et al. [4] proving that SPR is guaranteed for the polynomial family if and only if it holds for a finite number of polynomials in the family.

The second characterization of the uncertainty is specified directly in root space, i.e., some bounds for the locations of the roots of $A(z^{-1})$ are given. This type of description appears to be more useful, since an imperfect knowledge of the plant dynamics will be easier to translate into a root uncertainty set. Indeed, in the worst case, since $A(z^{-1})$ is assumed to be stable, it is clear that the poles of the uncertain plant are known to be inside the unit circle. Following the work in [5] and [6], we have recently presented a class of uncertainty sets for the roots of $A(z^{-1})$ [7] which allows us to guarantee the SPRness of the uncertain family in terms of a finite number of members of the family. This was done by means of the so-called *convex phase arcs* and the allowable sets include, as a special case, rectangles and circles in which it is possible to embed the roots in a very straightforward way.

In the more general case, the denominator of the unknown plant satisfies $A(z^{-1}) \in \mathcal{A}(z^{-1})$, where $\mathcal{A}(z^{-1})$ is the set of possible polynomials. Of course, the amount of uncertainty will be reflected in the size of $\mathcal{A}(z^{-1})$. Clearly, if the following condition holds

$$\operatorname{Re}\{A(e^{-j\omega})\} > 0, \quad \forall \omega \in [0, 2\pi), \quad \forall A(z^{-1}) \in \mathcal{A}(z^{-1}) \quad (4)$$

there will be no need for the filter $C(z^{-1})$ since $A(z^{-1})$ will already be SPR. Unfortunately, for moderate degree polynomials, or those having roots close to the unit circle, (4) does not hold, so it is necessary to filter the output error. In any case, a necessary condition for the existence of such a filter in terms of the arguments of the members of $\mathcal{A}(z^{-1})$ was given in [4] as

$$\max_{A \in \mathcal{A}} \{\arg A(e^{-j\omega})\} - \min_{A \in \mathcal{A}} \{\arg A(e^{-j\omega})\} < \pi \quad (5)$$

for all $\omega \in [0, 2\pi)$ where \arg denotes the unwrapped phase. In those cases where a finite number of members of the family is sufficient, the maximization and minimization of the phase is carried out within a discrete set.

Note that a necessary condition for the SPRness of A/C is that the filter $C(z^{-1})$ be minimum phase. In fact, it was proven that (5) is not only necessary but a sufficient condition for the existence of a minimum phase $C(z^{-1})$ guaranteeing (3) for every possible $A(z^{-1})$ [4].

3. FILTER DESIGN

Note that the design of $C(z^{-1})$ can be performed off-line, based on the available specifications for the uncertainty. The first attempt to design such a filter was presented in [4] and although the procedure is not very practical, we will outline it here, since it contains several ideas that we will

exploit later. Let

$$\phi^-(\omega) = \min_{A \in \mathcal{A}} \{\arg A(e^{-j\omega})\} \quad (6)$$

$$\phi^+(\omega) = \max_{A \in \mathcal{A}} \{\arg A(e^{-j\omega})\} \quad (7)$$

Then, it is clear that a minimum-phase filter with phase

$$\phi_o(\omega) = [\phi^+(\omega) + \phi^-(\omega)]/2 \quad (8)$$

will make every possible $A(z^{-1})$ SPR, since

$$|\arg A(e^{-j\omega}) - \phi_o(\omega)| < \pi/2 \quad (9)$$

By means of the Hilbert transform it is possible to find a minimum phase $F(z^{-1})$, such that $\arg F(e^{-j\omega}) = \phi_o(\omega)$. Then, from the minimum-phase property of $F(e^{j\omega})$, it follows that a truncated Laurent series of $F(z^{-1})$ can be carried out to obtain an approximation $C_N(z^{-1})$ of degree N such that the SPR condition holds with this output filter. It is obvious that N has to be made large enough to guarantee that the filter is minimum phase and it maintains the SPR condition.

Another interesting approach has been recently proposed by Tesi et al. [6] and it deals with simple FIR filters that guarantee the SPRness of a specific class of uncertainties. Unfortunately, this class of allowable uncertainties is very restrictive and furthermore the condition (3) deals with regions of the complex plane other than the unit circle. Since this definition of SPRness is suitable for highly colored inputs, we will not consider this technique further although it should be mentioned that it produces low-order and very structured FIR filters.

Yet another design method has been presented in [8]. It is restricted to plant denominators of the interval class (a subclass of the polytopic family) and provides sufficient conditions for the existence of C of a certain degree together with a cumbersome design procedure. Moreover, it is only valid for continuous-time polynomials, and its extension to the discrete-time case is difficult.

4. DESIGNING WITH PHASE CONSTRAINTS

Our starting point is the paper by Anderson et al. [4], in which the importance of matching the phase of $\phi_o(j\omega)$ in (8) is recognized. However, this is not the only possible phase function, since depending on the uncertainty, there is some freedom in choosing $C(z^{-1})$. This freedom is important because we will be interested in minimizing the degree of $C(z^{-1})$. Let $\phi_C(\omega)$ denote the unwrapped phase of a minimum-phase $C(z^{-1})$, then a necessary and sufficient condition for robust SPRness is

$$\phi^+(\omega) - \pi/2 < \phi_C(\omega) < \phi^-(\omega) + \pi/2 \quad (10)$$

Note that this specifies a phase band that can be exploited to achieve lower-degree polynomials and can be used in the design procedure to determine a weighting function. An obvious fact is that if $C(z^{-1})$ is of degree N , then with N samples of $\phi_o(\omega)$ in (8) it is possible to obtain a filter whose phase interpolates these N samples. To confirm the

validity of this statement, consider that at every frequency ω_k , $k = 1, \dots, N$ we have

$$\alpha_k = \tan \phi_0(\omega_k) = \frac{-\sum_{n=1}^N c_n \sin n\omega_k}{1 + \sum_{n=1}^N c_n \cos n\omega_k} \quad (11)$$

Cross-multiplying in (11) we arrive to the following system of linear equations

$$\mathbf{S}\mathbf{c} + \alpha = 0 \quad (12)$$

with solution $\mathbf{c} = \mathbf{S}^{-1}\alpha$ [9]. The designer would then calculate \mathbf{c} , the vector of coefficients of $C(z^{-1})$ and then check if it meets the phase requirements in (10). It should be mentioned that when ϕ^+ and ϕ^- are obtained from a finite number of members of $\mathcal{A}(z^{-1})$, it is possible to efficiently evaluate (10) for every frequency with an exact and finite test. However, with this approach, we are not using the tolerance band given by (10) and, moreover, the frequency selection procedure is likely to be inadequate. To overcome these two problems, it is possible to use a weighted least-squares approximation (WLS) as described next: First, suppose that we have $L > N$ frequencies so in our attempt to solving each of equations (11) an error ϵ_k is made, so they can be rewritten as

$$\mathbf{S}\mathbf{c} + \alpha = \epsilon \quad (13)$$

where now \mathbf{S} is an $L \times N$ matrix. Suppose also that we have a weighting function, specified by the coefficients w_k , $k = 1, \dots, L$ so the weighted squared error

$$\sum_{i=1}^L w_i \epsilon_i^2 = \epsilon^T \mathbf{W} \epsilon \quad (14)$$

is minimized. Then, the optimal solution is

$$\mathbf{c} = (\mathbf{S}^T \mathbf{W} \mathbf{S})^{-1} \mathbf{S}^T \mathbf{W} \alpha \quad (15)$$

However, with this approach the equation error is minimized but not the error in the phase. In order to do that, consider that an error e_k is allowed in the phase $\phi_o(\omega_k)$. Then, equation (11) can be rewritten as

$$\frac{\sin(\phi_o(\omega_k) + e_k)}{\cos(\phi_o(\omega_k) + e_k)} = \frac{-\sum_{n=1}^N c_n \sin n\omega_k}{1 + \sum_{n=1}^N c_n \cos n\omega_k} \quad (16)$$

Expanding the left-hand term, cross-multiplying and dividing by $\sin \phi_o(\omega_k)$ we obtain

$$\sin \phi_o(\omega_k) X + \tan e_k \cos \phi_o(\omega_k) X = \cos \phi_o(\omega_k) Y + \tan e_k \sin \phi_o(\omega_k) Y$$

where X and Y are, respectively, the numerator and denominator of the right term of (16). Identifying terms with (13), we arrive to

$$\epsilon_k = [\cos \phi_o(\omega_k) X + \sin \phi_o(\omega_k) Y] \tan e_k \quad (17)$$

This allows us to obtain an adequate value for the weights. First, note that the absolute value of e_k will be bounded by $\Delta(\omega_k) = \phi^+(\omega_k) - \phi_o(\omega_k)$, so it seems reasonable to use

$$w_k = |(\cos \phi_o(\omega_k) X + \sin \phi_o(\omega_k) Y) \tan \Delta_k|^{-2} \quad (18)$$

Unfortunately, the weights depend on the actual value of X and Y which, in turn, depend on \mathbf{c} . Therefore, this can be used for devising an iterative procedure on which the weights are adjusted in terms of a tentative solution for \mathbf{c} and the problem solved using (15) for the weights matrix \mathbf{W} .

5. GUARANTEEING MINIMUM-PHASE

We will state first the following property that will prove to be useful: *The polynomial $C(z^{-1})$ with no zeros on the unit circle is minimum-phase if and only if its unwrapped argument $\phi_C(\omega)$ takes on the same values for $\omega = 0$ and $\omega = 2\pi$.* Therefore, since $\phi^+(\omega)$ and $\phi^-(\omega)$ are obtained from minimum-phase functions (all the members in $\mathcal{A}(z^{-1})$ are), it can be concluded that any FIR filter $C(z^{-1})$ satisfying the phase bounding condition (10) for **every** frequency in the interval $[0, 2\pi)$ will be minimum-phase. Therefore, it might be concluded that the minimum-phase condition is always satisfied. However, due to the finite number of interpolation points, it is difficult to interpolate with polynomials $C(z^{-1})$ satisfying the minimum-phase condition. The reason for this can be found in the inability of an arctangent-based formulation to track changes of phase that are integer multiples of π (on the other hand, this approach preserves the linearity of the equations). This implies that a polynomial with phase $\phi_C(\omega)$ interpolating exactly $\phi_o(\omega) + r\pi$ instead of $\phi_o(\omega)$, at a set of L given frequencies, will have zero squared-error, but it is not a valid solution. This problem may of course be overcome with a higher number of data points, at the expense of a significant increase in the computation cost, even if the FFT is used with equally spaced frequencies. In the remainder of this section we will concentrate on how to introduce constraints to guarantee the minimum-phase condition while additionally giving us a procedure for frequency selection.

Consider $C(z)$ instead of $C(z^{-1})$ by simply multiplying $C(z^{-1})$ by z^N , where N is the degree of $C(z^{-1})$, the phase condition in (10) transforms into

$$\varphi^-(\omega) < \varphi_C(\omega) < \varphi^+(\omega) \quad (19)$$

where $\varphi^-(\omega) = \phi^+(\omega) - \pi/2 + M\omega$, $\varphi^+(\omega) = \phi^-(\omega) + \pi/2 + M\omega$, $\varphi_C(\omega) = \arg C(e^{j\omega})$, and M is the difference of the degree of $C(z^{-1})$ and that of any of the $A(z^{-1})$. The minimum-phase condition can now be restated as: *The polynomial $C(z)$ with no zeros on the unit circle is minimum phase if and only if its unwrapped argument $\varphi_C(\omega)$ has a net change of $2N\pi$ as ω varies from $\omega = 0$ to $\omega = 2\pi$.* Again, satisfying (19) for every frequency in $[0, 2\pi)$ is necessary and sufficient for $C(z)$ of degree N to be minimum-phase. The question is whether this can be ensured by imposing a constraint at a finite number of frequencies. By the principle of the argument [10], $\varphi_C(\omega)$ has to be monotonously increasing for $C(z)$ to be minimum-phase. This idea can be exploited to force $\varphi_C(\omega)$ to increase by less than π from one frequency to the next [11], so it is ensured to be increasing. Therefore, we are interested in choosing $0 \leq \omega_1 < \omega_2 < \dots, \omega_L \leq \pi$ such that

$$\varphi_C(\omega_{k+1}) - \varphi_C(\omega_k) < \pi; \quad k = 0, \dots, L-1 \quad (20)$$

Since φ_C is not a priori known, we can use the bounds in (19) to give the following procedure: Given ω_k , pick ω_{k+1} satisfying

$$\phi^-(\omega_{k+1}) - \phi^+(\omega_k) + M(\omega_{k+1} - \omega_k) < 0 \quad (21)$$

and

$$\varphi^+(\omega_1) < \pi \quad (22)$$

$$\varphi^-(\omega_L) > \pi N - \pi \quad (23)$$

A careful look at (21) together with the monotonicity of φ^+ and φ^- reveals that the larger $\varphi^+ - \varphi^-$ is, the closer the consecutive frequencies have to be, which obviously implies a larger L . A consequence of this is that as the band in (10) is reduced in width (more uncertainty), less frequencies are needed to guarantee minimum-phase. This may be exploited to deliberately introduce ‘‘uncertainty’’ in order to reduce the number of frequencies. On the other hand, the narrower the band, the more restricted the interpolation, which may imply the need for more frequencies. In conclusion, we have to trade the number of frequencies and the interpolation at critical frequencies.

Instead of the L_2 design in the previous section, equation (19) gives the basis for an L_∞ interpolation. Since $\varphi_C(\omega_k)$ has to belong to the interval $(\varphi_k^-, \varphi_k^+)$, $k = 1, \dots, L$, this can be easily translated into a pair of linear inequalities for every ω_k involving the coefficients of $C(z^{-1})$. From $\varphi_C(\omega_k) < \varphi^+(\omega_k)$ we can conclude that

$$\text{Im}\{C(e^{j\omega_k})e^{-j\varphi^+(\omega_k)}\} > 0 \quad (24)$$

which is equivalent to

$$-\{(\cos(N\omega_k) + \sum_{n=1}^N c_n \cos[(N-n)\omega_k]) \sin(\varphi^+(\omega_k)) + \{\sin(N\omega_k) + \sum_{n=1}^N c_n \sin[(N-n)\omega_k]\} \cos(\varphi^+(\omega_k))\} > 0$$

and a similar inequality is obtained for $\varphi_C(\omega_k) > \varphi^-(\omega_k)$ substituting φ^+ by φ^- and > 0 by < 0 . By examining the nature of these two relations and considering the inequalities for the L different frequencies, it is possible to write

$$\mathbf{T}\mathbf{c} + \mathbf{d} > 0 \quad (25)$$

where \mathbf{T} is an $2L \times N$ matrix. The solution to this feasibility problem can be found by means of linear programming.

6. CONCLUSIONS AND FUTURE WORK

We have presented a method for guaranteeing the convergence of a popular IIR adaptive algorithm, by designing a system that filters the output identification error. It turns out that this design has to be done entirely in the phase domain, but an important problem appears when the necessary condition of minimum-phase in the filter has to be guaranteed. We have shown how the design procedure can be carried out, mainly in the mean-square sense, although we have paved the way for a minimax design. By using some fundamental results of stability theory, we have been

able to give a procedure for computing the set of frequencies where the interpolation is specified. In the future, we will investigate the possibility of making the frequencies selection adaptive, while guaranteeing the minimum-phase conditions.

These results are currently being investigated in designing an adaptive IIR filter in a digital communications context. The scheme that we are currently developing identifies the channel response (assumed to be a rational function) so that the resulting estimate can be used either by copying the weights to construct an appropriate equalizer or to simplify the equivalent channel so that a Viterbi algorithm can be used. This work will be the topic of a future paper.

7. REFERENCES

- [1] J. J. Shynk, ‘‘Adaptive IIR filtering,’’ *IEEE ASSP Magazine*, pp. 4–21, April 1989.
- [2] K. Narendra and J. Taylor, *Frequency Domain Criteria for Absolute Stability*. Academic Press, 1973.
- [3] C. Johnson, ‘‘Adaptive IIR filtering: Current results and open issues,’’ *IEEE Trans. on Information Theory*, vol. IT-30, pp. 237–250, March 1984.
- [4] B. Anderson, S. Dasgupta, P. Khargonekar, F. Kraus, and M. Mansour, ‘‘Robust strict positive realness: Characterization and construction,’’ *IEEE Trans on Circ. and Syst.*, vol. CAS-37, pp. 869–876, 1990.
- [5] A. Tesi, G. Zappa, and A. Vicino, ‘‘Enhancing strict positive realness condition on families of polynomials by filter design,’’ *IEEE Trans on Circ. and Syst-I: Fund. Theory and Appl.*, vol. 40, pp. 21–32, January 1993.
- [6] A. Tesi, A. Vicino, and G. Zappa, ‘‘Design criteria for robust strict positive realness in adaptive schemes,’’ *Automatica*, vol. 30, no. 4, pp. 643–654, 1994.
- [7] F. Pérez and C. Abdallah, ‘‘Phase-convex arcs in root space and its application to robust SPR problems.’’ Submitted, 1994.
- [8] A. Betser and E. Zeheb, ‘‘Design of robust strictly positive real transfer functions,’’ *IEEE Trans. Circ. and Syst: I-Fund. Theory and Appl.*, vol. 40, pp. 573–580, September 1993.
- [9] M. Hayes, J. Lim, and A. Oppenheim, ‘‘Signal reconstruction from phase or magnitude,’’ *IEEE Trans. on Acoust. Speech and Sig. Proc.*, vol. 28, pp. 672–680, December 1980.
- [10] B. Barmish, *New Tools for Robustness Analysis*. Mac Millan, 1994.
- [11] R. Kaminsky and T. Djaferis, ‘‘A novel approach to the analysis and synthesis of controllers for parametrically uncertain systems,’’ *IEEE Trans on Auto. Control*, vol. 39, pp. 874–876, April 1994.