HYPERSTABLE POLYPHASE ADAPTIVE IIR FILTERS

C. Mosquera and F. Pérez-González

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

R. López-Valcarce

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

ABSTRACT

This work considers the implementation of recursive identification algorithms based on hyperstability concepts with polyphase structures. It is shown that the SPR condition required for convergence of these schemes can always be met by using a sufficiently high polyphase expansion factor M. For a given M, the degree of persistent excitation required for parameter convergence is obtained. When *a priori* knowledge about the unknown system is available, a compensating filter can be designed to avoid the need for a high M.

1. INTRODUCTION

Since the strictly positive real (SPR) character of the denominator of the unknown system, a requirement for convergence of hyperstable adaptive IIR filtering algorithms [1, 2], need not be fulfilled in practice, many suggestions have been made in the last years in order to avoid this drawback. In [3] the use of an overparameterized model was suggested, claiming that a sufficiently high degree of overparameterization would ensure convergence of the output error to zero. However, convergence of the parameter vector need not take place, due to the non-uniqueness in the representation of the system. On the other hand, if pole-zero cancellations are still used (in order to force the SPR condition) but in such a way that the model obtained after convergence is unique, then the previous problem disappears. This can be done by using a polyphase structure for the filter. Such structures have been proposed in [4] to improve convergence speed of gradient-based adaptive IIR filters. We are concerned here about the SPR condition and persistent excitation requirements that must apply when the polyphase form is used for hyperstable algorithms: it will be shown that the former can always be fulfilled for an appropriate choice of the polyphase expansion factor (an inappropriate choice may yield a non-SPR polynomial even if the original polynomial was SPR). The price to pay is a linear increase in the persistent excitation degree of the input signal. Also, for very high expansion factors, convergence speed decreases due to the existence of many adaptive parameters. If some *a priori* knowledge about the pole locations of the plant is available, the expansion factor can be reduced by properly designing a compensating filter.

2. ALGORITHM WITH POLYPHASE STRUCTURE

Consider a plant input-output equation

$$y(n) = H(z)u(n) = \frac{B_*(z)}{A_*(z)}u(n),$$

where $B_*(z)$ and $A_*(z)$ are coprime polynomials of degree N in z^{-1}. Now, if $p_i, i = 1, 2, \ldots N$ are the roots of $A_*(z)$, define the polynomial

$$P(z) = \prod_{i=1}^{N} \prod_{k=1}^{M-1} (1 - p_i e^{\mathrm{j}\frac{2\pi k}{M}} z^{-1}),$$

where M is the polyphase expansion factor. We can introduce $N(M-1)$ pole-zero cancellations in the plant model by means of $P(z)$ to obtain

$$H(z) = \frac{B_*(z)P(z)}{A_*(z)P(z)} = \frac{F_*(z)}{D_*(z^M)}, \tag{1}$$

with $F_*(z) = B_*(z)P(z)$ a polynomial of degree NM and

$$D_*(z) = 1 + d_1^* z^{-1} + \cdots + d_N^* z^{-N}.$$

The representation of $H(z)$ as in (1) is known as the M-fold polyphase form. Note that this form is unique since the factor $P(z)$ is uniquely determined by $H(z)$.

With this structure for the plant, one can use a similar form for the adaptive model:

$$\hat{H}(z) = \frac{F(z)}{D(z^M)} = \frac{\sum_{i=0}^{NM} f_i z^{-i}}{1 + \sum_{j=1}^{N} d_j z^{-jM}},$$

and now an adaptive identification algorithm can be developed. Define the parameter vector θ_n and the regressor vector X_n (both of size $(M+1)N+1$) as follows, with i ranging from 0 to NM and j from 1 to N:

$$\begin{aligned}\theta_n &= [\ f_i(n) \ \mid\ -d_j(n)\ ]^t, \\ X_n &= [\ u(n-i)\ \mid\ x(n-jM)\ ]^t,\end{aligned}$$

where $x(n)$ is the *a posteriori* estimate given by $x(n) = \theta_{n+1}^t X_n$. The *a posteriori* error is

$$e_o(n) = [y(n) - x(n)] + \sum_{k=1}^{R} c_k [y(n-k) - x(n-k)],$$

where $c_1, \ldots c_R$ are suitable constants. The corresponding *a priori* quantities are the estimate $\hat{y}(n) = \theta_n^t X_n$ and the error

$$e(n) = [y(n) - \hat{y}(n)] + \sum_{k=1}^{R} c_k [y(n-k) - x(n-k)].$$

With this, the adaptive algorithm can be written as

$$\theta_{n+1} = \theta_n + \frac{\mu X_n e(n)}{1 + \mu X_n^t X_n}, \tag{2}$$

with $\mu > 0$ a suitable stepsize.

Lemma 1 *Let* $C(z) = 1 + \sum_{k=1}^{R} c_k z^{-k}$. *If the transfer function* $C(z)/D_*(z^M)$ *is SPR, i.e. if the system* $C(z)/D_*(z^M)$ *is stable and causal and*

$$\mathrm{Re}\, \frac{C(e^{j\omega})}{D_*(e^{j\omega M})} > 0 \ \forall\ \omega,$$

then the algorithm (2) is asymptotically stable, that is, $e_o(n) \to 0$ *as* $n \to \infty$.

The proof can be found in [5]. Note that the compensating filter $C(z)$, which is chosen by the designer so that $C(z)/D_*(z^M)$ is SPR, can be IIR, i.e. $R = \infty$.

The important point is that the transfer function that should be made SPR is $C(z)/D_*(z^M)$. Assume for the moment that $C(z) = 1$. The SPR condition on $1/D_*(z^M)$ is equivalent to the SPR condition on $1/D_*(z)$. Note that the roots of $D_*(z)$ are p_i^M; if $H(z)$ is stable, then $|p_i| < 1$ and as M increases, $|p_i^M| \to 0$. This means that for high M, $D_*(z)$ can be made SPR. For example, if $H(z)$ is the discrete-time transfer function obtained by oversampling a continuous-time system, the roots p_i will tend to cluster around $z = 1$ in the complex plane. Under these conditions, $1/A_*(z)$ is very unlikely to be SPR. The transformation $p_i \to p_i^M$ tends to pull the roots away from this region.

Although the hyperstability theorem guarantees convergence of the output error $e_o(n)$ to zero if the SPR requirement is met, it does not guarantee parameter convergence to the true vector θ_*. For this to hold, persistent excitation (PE) conditions on the input signal are needed, which also gives exponential convergence of the adaptive algorithm [6]. In the standard algorithm [i.e. (2) with $M = 1$], this is satisfied if the power spectral density of the input signal u is nonzero at least at $2N+1$ different frequencies in $[0, 2\pi)$ [6]. For the polyphase structure, the number of nonzero frequencies in u must be at least $(M+1)N+1$ (the number of parameters in the adaptive model). This can be obtained using the same arguments as in [6]; see also [5]. This number increases linearly with M, so the price to pay for alleviating the SPR condition is a stronger PE requirement in the input signal.

3. SIMPLIFIED ALGORITHM

The algorithm above converges globally (under fulfillment of the SPR and PE conditions) for all $\mu > 0$. However, the plant output y was assumed to be noise free. Usually this is not the case, and μ has to be kept small in order to cope with the noise in e_o. Under this 'slow adaptation' the algorithm can be simplified [1, 7]:

$$\theta_{n+1} = \theta_n + \mu X_n e(n), \tag{3}$$

with X_n now simply given by

$$X_n = [\ u(n-i)\ \mid\ \hat{y}(n-jM)\ ]^t.$$

As before, $\hat{y}(n) = \theta_n^t X_n$, and $e(n) = C(z)[y(n) - \hat{y}(n)]$.

Again, it can be easily verified that for a fixed parameter vector θ,

$$e(n) = \frac{C(z)}{D_*(z^M)} s(n) + C(z)\eta(n),$$

with $s(n) = \tilde{\theta}_n^t X_n$, $\tilde{\theta}_n = \theta_* - \theta_n$ and $\eta(n)$ the noise component in the plant output. For slow adaptation, we can link the convergence properties of the algorithm to those of an associated ordinary differential equation (ODE) [2]; assuming that η is zero-mean and uncorrelated with u, for the algorithm (3) this ODE turns out to be

$$\frac{d\tilde{\theta}}{dt} = -\mathbf{R}\tilde{\theta},$$

with

$$\mathbf{R} = E\left[X_n \cdot \frac{C(z)}{D_*(z^M)} X_n^t\right].$$

The signals in $\mathbf{R}$ are functions of $\tilde{\theta}$ and therefore so is $\mathbf{R}$ itself.

Lemma 2 *If* $C(z)/D_*(z^M)$ *is SPR, the input signal* u *is persistently exciting of degree at least* $2MN+1$, *and* $\hat{H}(z)$ *has degree* NM *or* N, *then* $\mathbf{R}$ *is positive definite.*

See [5] for a proof. Thus under these conditions, the ODE associated to the simplified algorithm is globally convergent.

4. ROBUST DESIGN

As we have seen, the SPR condition for the algorithms (2) and (3) is given on the transfer function $C(z)/D_*(z^M)$, with $D_*(z^M) = \prod_{i=1}^{N}(1 - p_i^M z^{-M})$. Suppose that $C(z) = 1$ and that $\rho < 1$ is an upper bound on the magnitude of the plant poles. Then a sufficiently high M will make $D_*(z^M)$ SPR [5]:

$$M > \frac{\log(\sin\frac{\pi}{2N})}{\log\rho}. \quad (4)$$

The use of an appropriate compensator $C(z)$ may relax the bound on M. If some *a priori* knowledge about the pole locations of the plant is available, we can try the design of $C(z)$ such that $C(z)/A_*(z)$ is SPR, for $M = 1$. Thus, we consider an uncertain set $\mathcal{A}(z)$ for $A_*(z)$, built by all the polynomials of degree N with all their roots in some uncertainty region Ω inside the unit circle. The robust SPR problem, presented in [8], is defined as finding $C(z)$ such that $C(z)/A_*(z)$ is SPR for all $A_*(z) \in \mathcal{A}(z)$. For certain types of regions Ω, the phase of $A_*(z) \in \mathcal{A}(z)$, for z on the unit circle, is bounded above and below by the phase of a finite number of polynomials, $A_1(z), \cdots, A_r(z)$, known as extreme polynomials. Thus, the search for the solution of the robust SPR problem, $C(z)$, is reduced to finding $C(z)$ such that $C(z)/A_1(z), \cdots, C(z)/A_r(z)$ are SPR. The necessary and sufficient condition for the existence of such $C(z)$ was presented in [8]: There exists $C(z)$ such that $C(z)/A_1(z), \cdots, C(z)/A_r(z)$ are SPR if and only if for all ω in $[0, 2\pi)$

$$\max_i[\arg\{A_i(e^{j\omega})\}] - \min_i[\arg\{A_i(e^{j\omega})\}] < \pi. \quad (5)$$

As an example, if Ω is a circle centered at c real, and with radius ρ, then it has two extreme polynomials, namely, $A_1(z) = (1-(c-\rho)z^{-1})^N$ and $A_2(z) = (1-(c+\rho)z^{-1})^N$ [10]. If $M > 1$, we can design a transfer function $G(z)$ to make the uncertain set $\mathcal{D}(z)$ SPR, where

$$\mathcal{D}(z) = \left\{D_*(z), D_*(z) = \prod_{i=1}^{N}(1 - p_i^M z^{-1}), p_i \in \Omega\right\}.$$

By Lemma 1, and considering again that a transfer function $P(z^M)$ is SPR if and only if $P(z)$ is SPR, $C(z) = G(z^M)$ must be the compensator used in the algorithm, with $G(z)$ such that $G(z)/D_*(z)$ is SPR for all $D_*(z)$ in $\mathcal{D}(z)$. For example, if Ω denotes a circle centered at the origin with radius ρ, the bound on M such that the family $\mathcal{D}(z)$ verifies (5) is [5]

$$M > \frac{\log(\tan\frac{\pi}{2N})}{\log\rho}, \quad (6)$$

lower than the bound in (4).

Next we present a procedure to design the appropriate compensator $C(z)$ for uncertainty sets $\mathcal{D}(z^M)$, where $\mathcal{D}(z)$ can be described by two extreme polynomials $D_1(z)$ and $D_2(z)$, and M denotes the polyphase factor. Notice that $\mathcal{D}(z) = \mathcal{A}(z)$ for $M = 1$. First we design $G(z)$ such that $G(z)/D_1(z)$ and $G(z)/D_2(z)$ are simultaneously SPR. If we write

$$\begin{aligned} Q(z) &= D_1(z)G(z^{-1}) + D_1(z^{-1})G(z), \\ R(z) &= D_2(z)G(z^{-1}) + D_2(z^{-1})G(z), \end{aligned}$$

after some straightforward computations $G(z)$ can be expressed in terms of $R(z)$ and $Q(z)$ as

$$G(z) = \frac{R(z)D_1(z) - Q(z)D_2(z)}{T(z)},$$

with $T(z) = D_1(z)D_2(z^{-1}) - D_2(z)D_1(z^{-1})$. The sign of $Q(z)$ is the same as the sign of $\mathrm{Re}\left\{\frac{G(z)}{D_1(z)}\right\}$ on the unit circle; an identical consideration can be made for $R(z)$ and $D_2(z)$. Therefore, $G(z)$ can be computed by finding $Q(z) = Q(z^{-1})$ and $R(z) = R(z^{-1})$ positive on the unit circle and such that

$$R(z)D_1(z) - Q(z)D_2(z) = 0$$

at the roots of $T(z)$ on the unit circle and outside the unit circle. The roots of $T(z)$ outside the unit circle can be readily included in $R(z)$ and $Q(z)$ (and by symmetry, the roots inside the unit circle). However, given the positivity of $R(e^{j\omega})$ and $Q(e^{j\omega})$ for all ω, the zeros of $T(z)$ on the unit circle must be canceled out by solving the following interpolation problem:

$$\frac{R(z)}{Q(z)} = \frac{D_2(z)}{D_1(z)}$$

at the roots of $T(z)$ on the unit circle. With $D_1(z) = (1 - \rho^M z^{-1})^N$ and $D_2(z) = (1 + \rho^M z^{-1})^N$, those roots are only two, namely, $z = 1$ and $z = -1$. The interpolation can be carried out by making use of the algorithm presented in [9], and the degree of $C(z) = G(z^M)$ will be NM [5].

5. SIMULATIONS

Let us consider an identification setting with a second-order plant and with a colored input $u(n)$ taken as the output of the filter $S(z) = 1/(1 - 1.2z^{-1} + 0.7z^{-2})$, when driven with zero mean, unit variance white noise. The poles are known to lie in a disk Ω centered at the origin, with radius 0.95. According to (4), the smallest M ensuring convergence for any set of poles in Ω is 7, if $C(z) = 1$. However, for $M = 1$, we can obtain a robust compensator $C(z)$, following the steps exposed above: $C(z) = 1 + 0.0974z^{-1} - 0.9025z^{-2}$. The true plant considered was $H(z) = (1 + 2z^{-1} + z^{-2})/(1 - 1.8z^{-1} + 0.9z^{-2})$.

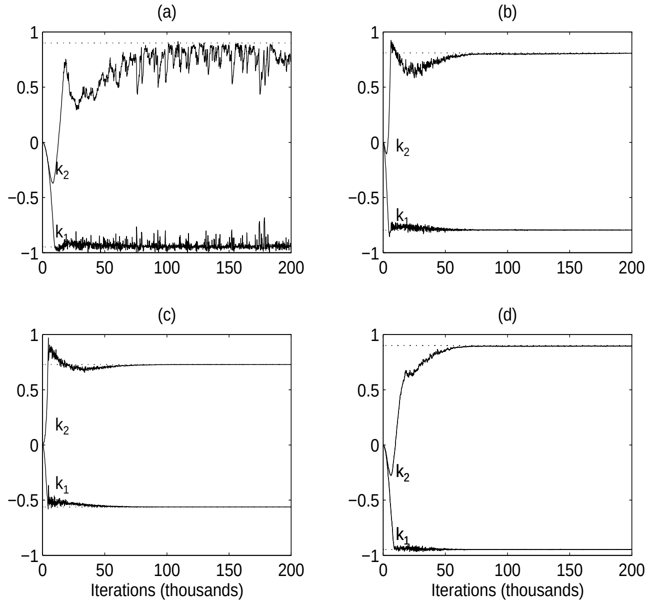

Figure 1: Trajectories of the reflection coefficients of the adaptive filter. (a) $M = 1, C(z) = 1$, (b) $M = 2, C(z) = 1$, (c) $M = 3$, $C(z) = 1$, and (d) $M = 1$, $C(z) = 1 + 0.0974z^{-1} - 0.9025z^{-2}$. The stepsizes for the recursive and non-recursive part were 10^{-6} and 10^{-5} respectively.

The simplified version of the algorithm was used, with the recursive part implemented as a two-multiplier lattice to allow for stability monitoring, following the approach of [11] (Note that the reflection coefficients of $D(z^M)$ are all zero except $k_M, k_{2M}, \cdots, k_{NM}$). Figure 1 shows the results of the polyphase implementation for several values of M with $C(z) = 1$. The compensator computed above is also implemented with $M = 1$. For $M = 1$, $C(z) = 1$, ill-convergence is observed, due to the fact that the plant is not SPR. For $M = 2$, $C(z) = 1$, the corresponding $D_*(z)$ is not SPR yet; however, the range of frequencies in which its real part is negative is smaller than that of $A_*(z)$ and convergence is achieved anyway. For $M = 3$, $C(z) = 1$ and $M = 1$, $C(z) = 1 + 0.0974z^{-1} - 0.9025z^{-2}$ the corresponding transfer functions are SPR and convergence is guaranteed for any persistently exciting input signal.

6. CONCLUSIONS

An analysis of hyperstable polyphase adaptive IIR filters has been performed. By appropriately choosing the polyphase expansion factor M, the SPR condition required for convergence can be satisfied. M can be selected using *a priori* information about the unknown plant poles in the form of uncertainty regions. Two drawbacks of the polyphase representation are a stronger PE requirement in the input signal and a higher number of adaptive parameters. The former can be overcome if the designer is able to choose the input signal. The latter can be considerably alleviated through the design of an appropriate compensating filter, with a procedure which uses *a priori* information.

7. REFERENCES

[1] C. R. Johnson, Jr, *Lectures on Adaptive Parameter Estimation*. Prentice Hall, 1988.

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

[3] S. Shah and G.F. Franklin. "On satisfying strict positive real condition for convergence by overparameterization", IEEE Trans. on Automatic Control, vol. 27, pp. 715–716, June 1982.

[4] P.M.S. Burt and M. Gerken. "A polyphase IIR adaptive filter: Error surface analysis and application", Proc. International Conference on Acoustics, Speech and Signal Processing, pp. 2285–2288, April 1997.

[5] R. López-Valcarce, C. Mosquera and F. Pérez-González, "Hyperstable adaptive IIR algorithms with polyphase structures: analysis and design", IEEE Trans. on Signal Processing, accepted for publication.

[6] B.D.O. Anderson and C.R. Johnson, Jr, "Exponential convergence of adaptive identification and control algorithms", Automatica, vol. 18, pp. 1–13, May 1982.

[7] M.G. Larimore, J.R. Treichler and C.R. Johnson, Jr., "SHARF: an algorithm for adapting IIR digital filters", IEEE Trans. on Acoustics, Speech and Signal Processing, vol. 28, pp. 428–440, August 1980.

[8] B.D.O. Anderson, S. Dasgupta, P. Khargonekar, F.J. Kraus, and M.Mansour, "Robust strict positive realness: Characterization and construction", IEEE Trans. on Circuits and Systems, vol. 37, pp. 869–876, July 1990.

[9] C. Mosquera and F. Pérez, "An algorithm for interpolation with positive rational functions on the imaginary axis", Automatica, vol. 33, pp. 2277–2280, December 1997.

[10] A. Tesi, A. Vicino and G. Zappa, "Design criteria for robust strict positive realness in adaptive schemes", Automatica, vol. 30, pp. 643–654, April 1994.

[11] K.X. Miao, H. Fan and M. Doroslovački, "Cascade normalized lattice adaptive IIR filters", IEEE Trans. on Signal Processing, vol. 42, pp. 721–42, April 1994.