STRUCTURAL ISSUES IN CASCADE-FORM ADAPTIVE IIR FILTERS

Geoffrey A. Williamson
Dept. of Elec. & Comp. Engr.
Illinois Inst. of Tech.
Chicago, IL 60616

James P. Ashley
Cellular Infrastructure Group
Motorola, Inc.
Arlington Heights, IL 60004

Majid Nayeri
Dept. of Elec. Engr.
Michigan State Univ.
East Lansing, MI 48824

ABSTRACT

Adaptive IIR filters implemented in cascade-form are attractive due to the ease with which their stability may be monitored. In this paper, four cascade-form structures are compared for use in adaptive filtering with respect to complexity of implementation, error surface geometry, and adaptation speed. The four structures include a cascade of second order pole/zero sections, a cascade of second order all-pole sections followed by a tapped delay line, and two new structures introduced by this paper. The latter pair includes a tapped cascade, which is a cascade of second order all-pole sections whose output is constructed as a weighted combination of signals tapped from the cascade. The second new structure is a modification of the tapped cascade that yields orthogonal signals at the taps of the cascade. It is shown that the tapped cascade provides the best overall performance in the respects noted above.

1. INTRODUCTION

A number of adaptive signal processing applications, such as acoustic echo cancellation, demand long impulse responses for adequate performance of the adaptive filter. This fact motivates examination of adaptive IIR filters, which can provide long impulse responses with a computational burden much reduced in comparison with adaptive FIR filters. Cascade-forms provide an attractive realization for adaptive IIR filters because stability of the filter parametrization is easily monitored, and because filter pole locations are readily obtained from the adapted parameters. Recent papers, such as [4, 5, 8] among others, have reflected interest in this type of adaptive filter. In this paper, we examine different implementation structures for cascade-form adaptive IIR filters with respect to the complexity of the sensitivity function generation for gradient descent adaptation algorithms, the geometry of the error surface for the adaptive filter, and the convergence speed obtained in adaptation.

2. REALIZATIONS OF CASCADE-FORM ADAPTIVE FILTERS

The typical cascade-form filter is implemented as a cascade of second-order sections, each having two zeros and two poles, as shown in Fig. 1, together with an overall gain parameter b_0. This structure was considered for adaptation in [4, 8]. We shall refer to this form as the *standard cascade*, or SC.

A different realization, considered in [9], realizes the filter numerator as an all-zero tapped delay line section, with an all-pole cascade of second order sections following in series. In [5], the reverse is suggested: an all-pole cascade followed by a tapped delay line to realize the numerator, as shown in Fig. 2. We will refer to these as *all-pole-based cascades*, or APC.

In this paper, we introduce two new cascade-form realizations. The first of these structures has a cascaded all-pole section, with the filter output formed from weighted taps taken from the interior of the structure, as depicted in Fig. 3. We refer to it as a *tapped cascade*, or TC. It is shown in [13] that this structure is able to represent an arbitrary filter transfer function of a given order.

The second new structure, motivated by the Kautz network discussed in [10], augments the all-pole part of the tapped cascade with zeros to create orthogonal signals at the taps. We depict this *orthogonal tapped cascade* (OTC) in Fig. 4. Using the results given in [10], one may show that the tapped signals are orthogonal when the input is white.

3. ADAPTIVE ALGORITHMS

For each of these structures, we will consider use of the following two algorithms. Let $e(k) = d(k) - y(k)$ be the error signal obtained as the difference between the adaptive filter output $y(k)$ and the desired output signal $d(k)$. With $\theta(k)$ a vector of adaptive filter parameters, and with

$$\psi(k) = \frac{\partial y(k)}{\partial \theta(k)} \qquad (1)$$

the sensitivity function of the output with respect to those parameters, the normalized Least Mean-Square (NLMS) algorithm is given by

$$\theta(k+1) = \theta(k) + \frac{\mu}{\epsilon + \psi^T(k)\psi(k)}\psi(k)e(k). \qquad (2)$$

We use a normalized version of LMS to allow comparison of convergence for different structures for a given step size μ.

The Gauss-Newton (GN) algorithm is

$$P(k) = \frac{1}{\lambda}\left(P(k-1) - \frac{P(k-1)\psi(k)\psi^T(k)P(k-1)}{\lambda + \psi^T(k)P(k-1)\psi(k)}\right); \qquad (3)$$

$$\theta(k+1) = \theta(k) + P(k)\psi(k)e(k). \qquad (4)$$

Specification of a fixed forgetting factor λ enables comparison of convergence of the different structures when using the GN algorithm.

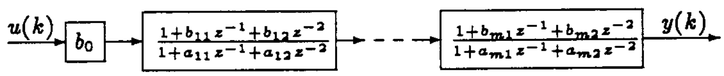

Figure 1: Standard cascade-form filter.

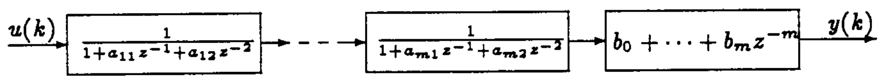

Figure 2: All-pole-based cascade-form filter.

4. COMPLEXITY OF SENSITIVITY FUNCTION GENERATION

Gradient based adaptive algorithms such as the LMS and GN algorithms all require computation of the sensitivity function $\psi(k)$. One difficulty with cascade structures is that the sensitivity function generation can be computationally demanding. For instance, the algorithm obtained by applying the sensitivity generation approach in [12] to the SC has complexity that is quadratic in the filter order.

Work in [11] establishes that computation of $\psi(k)$ requires a minimum of N multipications beyond those required to construct the filter output $y(k)$, where N indicates the filter order, for a total of $3N+1$ multiplications in the implementation. Using the terminology of [11], we refer to adaptive IIR filters that achieve this lower bound as *lowest complexity* adaptive filters.

4.1. Standard Cascade

Rao suggests a method to simplify the sensitivity generation for the SC in [5], based on an approach introduced in [2]. However, the simplification requires inversion of the filter's dynamics, so that if zeros lie outside the unit circle, the sensitivity generation will be unstable. Currently, the authors are aware of no lowest complexity implementation of the SC adaptive IIR filter.

4.2. All-pole-based Cascade

The same simplified sensitivity generation of [5] applies to the APC arrangement, but without the stability problems that occur with the SC as long as the numerator section is implemented second. As noted in [11], this realization is of lowest complexity.

4.3. Tapped Cascades

Figure 5 shows a realization of the TC, including the sensitivity generation, that was derived in [1] using techniques described in [11]. This implementation possesses the lowest complexity property of only $3N+1$ multiplications.

For the OTC of Fig. 4, however, no simplified sensitivity generation is known, and its complexity remains on the order of N^2. An implementation of the sensitivity generation may be derived using the procedures of [12], and is given in [1].

5. ERROR SURFACE GEOMETRY

The issue of error surface geometry for adaptive IIR filters is a complex one. We restrict consideration to whether corresponding error surface properties for direct-form adaptive IIR filters, which have received some study (see, *e.g.*, [7] and, more recently, [3]), accrue for the cascade-forms considered here.

If the relationship between a cascade-form filter's parameters and those of the direct-form is diffeomorphic, then the qualitative features of the error surface of the one carries over to the other, and *vice versa* [4]. However, due to the ambiguity in permutations of the ordering of a cascade-form filter's poles, this will never be the case. Nonetheless, the existence of *locally* diffeomorphic relationships between the various parametrizations at almost all values of the parameters enables some conclusions to be drawn.

Two possibilities can prevent the existence of locally diffeomorphic mappings between the parametrizations. The first is non-existence of a locally invertible mapping between the parametrizations at a given parameter value. The second is the failure of the invertible mapping to be differentiable in one or both directions.

5.1. Standard Cascade

Nayeri and Jenkins show that when the gain parameter of the SC is fixed at a unit value, only at parameter values corresponding to repeated pole or zero locations is it possible for a local diffeomorphism to fail to exist [4]. This failure is due only to non-differentiability of the mapping in one direction. Genericly, these points will be saddle points of the error surface. The theoretical gradient at such points preserves the repeated poles or zeros, so that trajectories will remain on manifolds corresponding to repeated poles or zeros if they are initialized there.[1] The implication of these arguments is that, given an initialization of the SC without repeated poles, repeated poles or zeros will not occur under theoretical gradient descent, and the SC's adaptation possesses all the qualitative features of the direct-form's.

However, if the gain parameter is included, no invertible mapping between direct-form and SC parameters exists at points corresponding to systems whose transfer functions are strictly proper [8]. Loss of the mapping itself is more

[1] Gradient noise will in general move the adaptation off these manifolds.

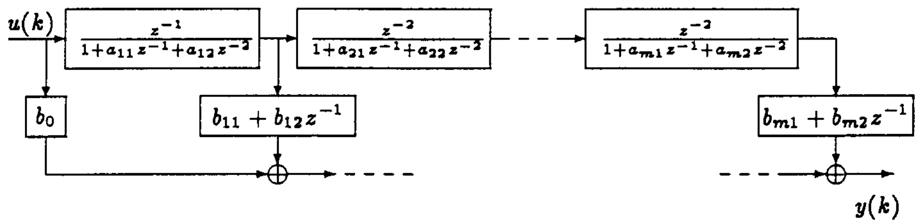

Figure 3: Tapped cascade-form filter.

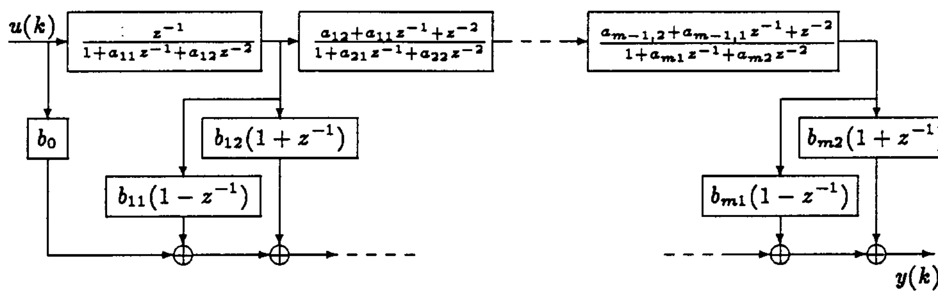

Figure 4: Orthogonal tapped cascade-form filter.

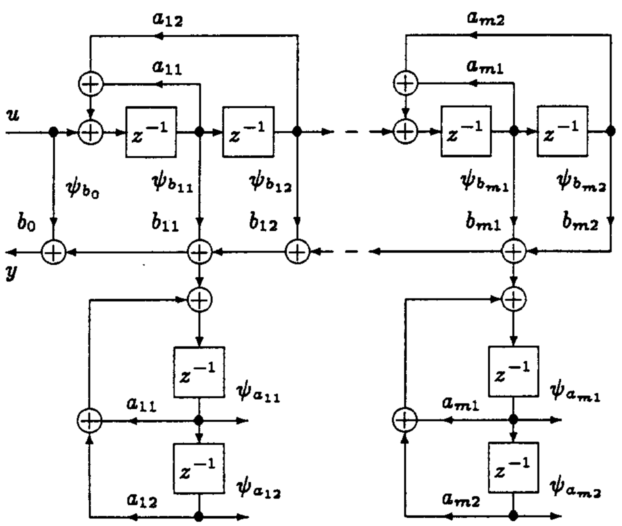

Figure 5: Tapped cascade realization and sensitivity generation.

catastrophic to the adaptation than simple failure of differentiability, as it is now possible for unbounded parameter growth to occur under the gradient descent, even in situations in which the direct-form adaptive IIR filter converges globally [8]. This clearly is undesirable adaptation behavior.

5.2. All-pole-based Cascades

One may show that the APC structure possesses a locally invertible mapping to the direct-form at all parameter values, and so its qualitative adaptation behavior is similar to the unit gain SC studied in [4]. In fact, the analysis of [4] applies directly to the APC, and it is somewhat simpler since the APC numerator parameters are the same as the direct-form's. This indicates that the all-pole-based adaptation will be qualitatively equivalent to a direct-form filter's adaptation.

5.3. Tapped Cascades

The tapped cascade structures, both TC and OTC, also possess locally invertible mappings to the direct-form at all parameter values. The mapping of the denominator parameters of the tapped cascades is identical to that for the other cascades, since all the cascade-forms factor the denominator into second order sections. Furthermore, the tapped cascades' outputs are linear in the numerator parameters; hence, for each set of fixed denominator (pole) parameters, there is a global minimum for the numerator parameters. Therefore, by the theory of [4], new saddle points in the tapped cascade error surface correspond only to parameter values reflecting repeated poles, and these will not alter the qualitative features of the error surface.

6. CONVERGENCE SPEED

To compare the convergence rates of the various cascade-form implementations, simulation experiments were conducted using both the NLMS algorithm (with $\mu = 0.4$ and $\epsilon = 0.001$) and the GN algorithm (with $\lambda = 0.95$ and $P(0) = 0.001I$). The stability check and projection scheme

Table 1: Convergence Statistics: NLMS Algorithm

	Convergence time (iterations $\times 10^3$) min	avg	max	Converge (%)	Unstable (%)
TC	1.20	46.4	226.4	51.4	0.0
OTC	1.00	50.5	230.0	51.2	0.0
APC	1.00	51.9	224.7	46.2	0.0
SC	3.00	60.1	205.2	30.2	0.0

Table 2: Convergence Statistics: GN Algorithm

	Convergence time (iterations $\times 10^3$) min	avg	max	Converge (%)	Unstable (%)
TC	0.37	1.15	16.2	61.8	0.0
OTC	0.36	1.72	21.8	55.4	0.0
APC	0.36	1.04	7.38	31.2	0.0
SC	0.39	2.00	23.4	27.4	0.4

for the pole locations as given in [6] was employed in all cases.

The desired signal $d(k)$ was the output of a randomly determined fourth order transfer function, normalized to unit power given a unit variance white Gaussian input. White Gaussian noise of variance 10^{-8} was added in $d(k)$. Table 1 shows the convergence statistics for 500 separate random systems using the NLMS algorithm. Each trial was conducted for 250,000 iterations, and each cascade structure was adapted using the same input and noise sequence. Table 2 shows the convergence statistics using the GN algorithm, but running for only 25,000 iterations. Convergence time is determined by the time in which the lowpass filtered squared output error falls and remains below an empirical threshold of 10^{-6}, a factor of 100 greater than the variance of the additive noise in the desired signal. Also indicated in the Tables is the percentage of trials in which the convergence conditions were met before the end of the trial, and the percentage of trials in which the output error rose above 10^9, essentially indicating a loss of stability. Trials which failed to show convergence are not reflected in the convergence time statistics.

It is apparent from the tables that the TC has the best overall performance.

7. CONCLUSIONS

Our examination indicates the following comparison between the different cascade stuctures. The SC is poor in sensitivity complexity and error surface geometry, while the OTC is poor in sensitivity complexity. TC and APC are both fine in these respects. Convergence speed favors TC, then OTC, with APC and finally SC finishing the ordering in this aspect. Overall, the evidence supports the TC as the most promising of the structures for use in adaptive filtering.

8. REFERENCES

[1] J.P. Ashley, "Cascade form structures for adaptive infinite impulse response filters," M.S. thesis, Illinois Institute of Technology, December 1994.

[2] L.B. Jackson and S.L. Wood, "Linear prediction in cascade form," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 26, pp. 518-528, December 1978.

[3] M. Nayeri, "A mildly weaker sufficient condition in IIR adaptive filtering," *IEEE Trans. Signal Processing*, vol. 42, pp. 203-206, January 1994.

[4] M. Nayeri and W.K. Jenkins, "Alternate realizations of adaptive IIR filters and properties of their performance surfaces," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 485-496, April 1989.

[5] B.D. Rao, "Adaptive IIR filtering using cascade structures," *Proc. 27th Asilomar Conf. on Signals, Syst., and Computers*, Pacific Grove, CA, pp. 194-198, November 1993.

[6] J.J. Shynk, "Adaptive IIR filtering," *IEEE ASSP Magazine*, vol. 6, pp. 4-21, April 1989.

[7] T. Söderström and P. Stoica, "Some properties of the output error method," *Automatica*, vol. 13, pp. 1692-1716, December 1982.

[8] G.A. Williamson, "Gradient-descent adaptation of cascade-form adaptive filters," *Proc. 27th Asilomar Conf. on Signals, Syst., and Computers*, Pacific Grove, CA, pp. 1559-1563, November 1993.

[9] F.X.Y. Gao and W.M. Snelgrove, "An adaptive backpropagation cascade IIR filter," *IEEE Trans. Circuits Syst. II: Analog and Digital Sig. Proc.*, vol. 39, pp. 606-610, September 1992.

[10] B. Wahlberg, "System identification using Kautz models," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 1276-1282, June 1994.

[11] G.A. Williamson, "Implementations of adaptive IIR filters with lowest complexity," *IEEE Trans. Circuits Syst. II: Analog and Digital Sig. Proc.*, under review.

[12] G.A. Williamson, C.R. Johnson, Jr., and B.D.O. Anderson, "Locally robust identification of linear systems containing unknown gain elements with application to adapted IIR lattice models," *Automatica*, vol. 27, pp. 783-798, September 1991.

[13] G.A. Williamson and S. Zimmermann, "Globally convergent adaptive IIR filters based on fixed pole locations," *IEEE Trans. Signal Processing*, under review.