IETE Journal of Research

ISSN: 0377-2063 (Print) 0974-780X (Online) Journal homepage: http://www.tandfonline.com/loi/tijr20

Mean-Convergence behavior of Adaptive Identification Algorithms for Pole-Zero Systems

Ratnam V Raja Kumar & Chunduri B Rama Rao

To cite this article: Ratnam V Raja Kumar & Chunduri B Rama Rao (1996) Mean-Convergence behavior of Adaptive Identification Algorithms for Pole-Zero Systems, IETE Journal of Research, 42:1, 27-32, DOI: 10.1080/03772063.1996.11415893

To link to this article: http://dx.doi.org/10.1080/03772063.1996.11415893

Published online: 26 Mar 2015.

Journal of the IETE
Vol. 42, No. 1, January-February 1996, pp 27-32

Mean-Convergence behavior of Adaptive Identification Algorithms for Pole-Zero Systems

RATNAM V RAJA KUMAR
Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology
Kharagpur 721 302, India.
AND
CHUNDURI B RAMA RAO
Department of Electronics and Communication, Regional Engineering College, Warangal 506 004, India.

Adaptive methods for the estimation of unknown system parameters has the advantage of tracking time-varying systems. Identification algorithms for recursive systems produce nonquadratic performance functions. In such problems it is very difficult to estimate the nature of convergence in a stochastic frame work. Recently, it has been shown that the ensemble mean parameter updating equations of IIR adaptive algorithm can be represened by associated ordinary differential equations (ODEs). A method of solving the ODEs in order to analyze the mean-convergence behavior of these algorithms, given the mean description of the input in the form of power spectral density, has been presented recently. In this paper, this procedure is applied to study the convergence behavior of recursive adaptive algorithms applied for the identification of pole-zero systems. Effectiveness of this method is shown through analytical and simulation results.

Indexing terms: System identification, Adaptive, Pole-zero systems, ODE, Mean-convergence.

ADAPTIVE estimation of systems with time-varying characteristics is an important area of current research [1-7]. Most of the work carried out in this area is available under the system identification and adaptive filtering literature. IIR adaptive algorithms have several advantages over their FIR counterparts. In many situations it is possible to approximate even an all zero system of a large order by a pole-zero system of a much lower order. Further, an IIR filter may represent an optimum structure for pole-zero systems which are commonly encountered in practice. The main drawbacks of an IIR adaptive filter are the need for stability monitoring during adaptation and uncertainty in the convergence time for stochastic inputs which is a result of the nonquadratic performance function. The nature of convergence and convergence time may be different even for the same input signal when the noise sample set is different. In the case of IIR adaptive filters, while handling stochastic signals, it is very difficult to estimate the nature of convergence and convergence time. If the estimator dynamics is known, some idea regarding the stability and the convergence time required can be obtained. Therefore in such situations it is of interest to estimate the nature including the time of ensemble mean convergence, at least. In this direction, recently, a method of analyzing the ensemble mean convergence behavior of recursive adaptive filters given the mean description of the input is presented in [17-18]. The basis for this anlaysis is the ordinary differential equation (ODE) representation of the recursive adaptive filters [9].

Orginally, the ODE approach was presented by L Ljung for the convergence analysis of recursive identification algorithms [8]. It was shown that the asymptotic behavior of an identification algorithm can be obtained by solving the ordinary differential equations corresponding to the identification algorithm. The recursive identification algorithms use an adaptive step size which tends to zero as time tends to infinity. Whereas in an adaptive filter the step size is a constant so that the algorithm can track the variations in the input. Later, the ODE approach was extended to IIR adaptive filters [9-15]. In [9] it was shown that the ensemble mean behaviour of IIR adaptive filters can be represented by their corresponding OEDs. Earlier, we applied this approach to study the convergence behavior of constrained IIR adaptive filters [9-10]. The ODEs representing an IIR adaptive filtering algorithm are nonlinear and it is extremely difficult to obtain a general closed form expression for convergence time and nature of convergence. It is only possible to evaluate and find solutions for specific filtering problems using numerical procedures.

In this paper we applied this procedure to study the convergence behavior of recursive adaptive identification algorithms for pole-zero systems. The nature and time taken for convergence of general examples of systems is studied. Location of the system poles relative to the unit circle decided the convergence nature of the identification algorithms. Hence in this work we have studied the dependence of the convergence behavior and the time taken for convergence on algorithm parameters and location of poles.

Through analytical method which involves computation of the ODE solutions numerically, the convergence behavior can be studied. This analysis provides a means to obtain a good

Paper No 106-B; received 1994 December 12; revised 1995 November 8;

idea about the nature of parameter adjustment, stability and convergence time. Effectiveness of this method is shown through some analytical and simulation results obtained from system identification examples.

THE FORMULATION OF AN ADAPTIVE ESTIMATOR

System model

Here, we considered the adaptive system identification problem as shown in Fig 1. The estimation of the unknown system is done by its inverse model. Let us consider the system transfer function as

$$H(z) = \frac{C(z)}{A(z)} \quad (1)$$

where $C(z) = 1 + c_1 z^{-1} + \ldots + c_p z^{-p}$

and $A(z) = 1 + a_1 z^{-1} + \ldots + a_q z^{-q}$, z^{-1} is the unit delay operator. Here, w_t and w'_t are the system input noise and output measurement error which are uncorrelated with each other. The time variable t is an integer which is a discrete time index. The SNR represents the ratio of the system output power to the additive measurement error power. The system output which is corrupted with the measurement noise is fed to the inverse model and the model parameters are estimated by whitening the output. The inverse model can also be viewed as a prediction error filter whose transfer function, input and output can be given as

$$G(z) = \frac{A(z)}{C(z)} \quad (2)$$

$$x_t = \text{System output + additive noise.} \quad (3)$$

$$\varepsilon_t = x_t - \phi_t^T \theta_{t-1} \quad (4)$$

where $\phi_t = [-x_{t-1} \ldots -x_{t-q} \varepsilon_{t-1} \ldots \varepsilon_{t-p}]^T$ at the time index t

$\theta_{t-1} = [a_1 a_2 \ldots a_q \; c_1 c_2 \ldots c_p]^T$ at time index $t-1$.

Algorithm recursions

A widely used criterion function for the estimation of the pole-zero system parameters is likelihood function of the conditional probability density function of the parameter vector given the observations. The RML algorithm is a recursive algorithm obtained for maximizing the likelihood function. The RML algorithm recursions for estimating the parameter vector $\hat{\theta}_t$ can be given as [3,4],

$$\varepsilon_t = x_t - \phi_t^T \hat{\theta}_{t-1} \quad (5a)$$

$$\hat{P}_t = \frac{1}{\lambda}\left[\hat{P}_{t-1} - \frac{\hat{P}_{t-1}\psi_t \psi_t^T \hat{P}_{t-1}}{\lambda + \psi_t^T \hat{P}_{t-1}\psi_t}\right] \quad (5b)$$

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \mu \hat{P}_t \psi_t \varepsilon_t \quad (5c)$$

where $\psi_t = \dfrac{\phi_t}{D(z)}$, $D(z) = 1 + \hat{k}\hat{c}_1 z^{-1} + \ldots + \hat{k}^p \hat{c}_p z^{-p}$ (5d)

Here λ is the forgetting factor ($0 < \lambda < 1$) used in computing the covariance matrix $\hat{P}_t$ recursively and μ is the adaptation step size. $\hat{k}$ is an algorithm parameter which controls the transient behavior of the algorithm [7]. Usually $\hat{k}$ is taken as 1. This algorithm is initialized by $\hat{P}(0) = \alpha I$ and $\hat{\theta}(0) = 0$,where α is a suitable scalar. In the Gauss-Newton algorithm form equation (5) can be expressed as [3-4].

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \mu \hat{R}_t^{-1} \psi_t \varepsilon_t \quad (6a)$$

$$\hat{R}_t = \lambda \hat{R}_{t-1} + \psi_t \psi_t^T \quad (6b)$$

where $\hat{R}_t = P_t^{-1}$, is proportional to the Hessian matrix of the G-N algorithm.

ODE REPRESENTATION

In [9], it is shown that the ensemble mean of the updating equations, (5*b*) and (5*c*) can be represented in the following ordinary differential equation form in continuous time, τ

$$\frac{d\theta_\tau}{d\tau} = \mu P_\tau E[\psi_t(\theta)\varepsilon_t(\theta)] \quad (7a)$$

$$\frac{dP_\tau}{d\tau} = \left(\frac{1}{\lambda} - 1\right) P_\tau - \frac{P_\tau E[\psi_t \psi_t^T] P_\tau}{\lambda + E[\psi_t^T P_{t-1} \psi_t]} \quad (7b)$$

Here, the denominator in (7*b*) involves computation of

$E[\psi_t^T P_{t-1} \psi_t]$

which can be expressed as $\sum_{i=1}^{q} \sum_{k=1}^{p} p_{ik} E[\psi_t(i)\, \psi_t(k)]$

where p_{ik} is the (i, k)th element of the matrix $[P]$ and $\psi_t(i)$ is the ith element of vector ψ_t.

For the ODE representation to be applicable, in addition to the regularity conditions required for ODE representation given in [8], the ensemble-mean representation requires an additional assumption on the input and the adaptive filter. It is necessary that the response of the adaptive filter attains steady state before the input charges appreciably. This means that there is a restriction on the rate at which the input signal

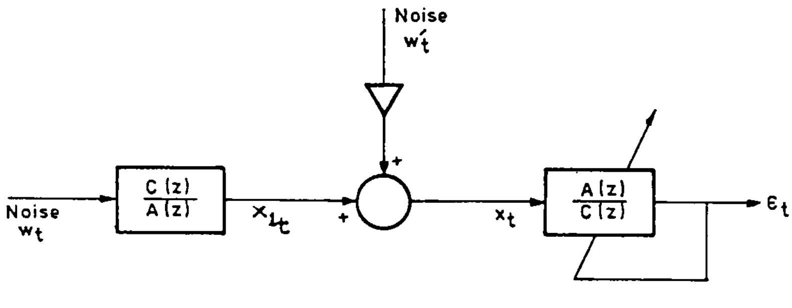

Fig 1 Block diagram of the system identification problem

characteristics can vary. That is, the rate of variation in the characteristics of the input should be much smaller than the inverse of the dominating time constant of the adaptive filter. Further in [9] it is shown that the solutions of the ODEs represent the mean trajectories of the parameter estimates, $\hat{\theta}_t$. It also means that the algorithms converge in mean to the stable stationary points of the ordinary differential equations.

As explained in the introduction, the mean convergence behavior of an adaptive filter in a specific application can be obtained by solving the ODEs corresponding to the algorithm, given the ensemble-mean description of the input such as the power spectral density (psd). When the input to be estimated is an ARMA process, the expectations of the gradient and the covariance matrix are nonlinear functions of θ_t. Then the differential equations are nonlinear by nature. Therefore, to obtain analytical solutions, given a general input description we have to resort to numerical procedures.

For solving these equations it is necessary to compute the expectations of $\psi(\hat{\theta}_t)\psi^T(\hat{\theta}_t)$, $\psi(\hat{\theta}_t)\,\varepsilon(\hat{\theta}_t)$ and $\psi(i,\hat{\theta}_t)\,\psi(k,\hat{\theta}_t)$ for $1 \le i \le q$, $1 \le k \le p$. Expressions for these are derived in [17,18], given the power spectral density of the input to system model, $\Phi_{xx}(z)$, as

$$E[\psi_t \psi_t^T] = \frac{1}{2\pi j}\oint \psi(z)\,\psi^+(z)\frac{dz}{z} \tag{8}$$

$$E[\psi_t\, \varepsilon_t] = \frac{1}{2\pi j}\oint \psi(z) \in (z)\,\frac{dz}{z} \tag{9}$$

and

$$E[\psi_t(i)\,\psi_t(k)] = \frac{1}{2\pi j}\oint \psi(i,z)\,\psi(k,z)\,\frac{dz}{z} \tag{10}$$

for $1 \le i \le q$, $1 \le k \le p$,

$$\text{where } \psi(z) = \frac{1}{D(z)}\,\Phi(z) \tag{11}$$

$$\text{and } \Phi(z) = [-z^{-1}X(z), \ldots, -z^{-q}X(z), z^{-1}G(z)X(z), \ldots, z^{-p}G(z)X(z)]^T \tag{12}$$

Here the + sign indicates the Hermitian operation.

These integrations can be computed using θ_t and $\Phi_{xx}(z)$ and help to solve differential equation (7).

The numerical procedure to obtain the ODE solutions can be summarized as follows. This numerical procedure is applied using a small time interval, $\Delta\tau$.

At any given time τ, given the parameter vector θ_τ and the input power spectral description $\Phi_{xx}(z)$, expectations of $\psi(\theta_\tau)\psi^T(\theta_\tau)$, $\psi(\theta_\tau)\varepsilon(\theta_\tau)$ and $\psi_\tau(i,\theta_\tau)\,\psi_\tau(k,\theta_\tau)$ can be obtained using eqns (8)-(10). Substituting these in the difference equations corresponding to ordinary differential eqns (7*a*) and (7*b*), the increments to the ODEs at time index *t* can be computed numerically. Using these increments $\theta_{\tau+\Delta\tau}$ can be obtained. Then using $\theta_{\tau+\Delta\tau}$ and input power spectral description, the same procedure can be repeated to get $\theta_{\tau+2\Delta\tau}$. Continuing this procedure until θ_τ attains a stage where θ_τ increments are negligible. That is, when θ_τ is very close to the optimum, the parameter trajectories can be obtained.

Accuracy of the numerical solutions depends on the time interval τ. A small time interval yields accurate results, but it requires a lot of computational time and *vice-versa*. A reasonable value for $\tau = 1$, the sampling time of the discrete version.

CONVERGENCE BEHAVIOURAL ANALYSIS

The parameter trajectories are the basis for the study of convergence behavior. Speed of convergence, dynamics of the estimates during the process of adaptation including the initial, intermediate and final convergence and the time taken for convergence *etc*, can be obtained from the ODE solutions. Dependence of these on the input SNR and the algorithm parameters also can be obtained from these characteristics. Although this analytical method can not provide closed form solutions for the convergence analysis, it provides us with ensemble convergence behavior. Individual convergence plots may deviate slightly from these characteristics. However, the deviation is not much and the ODE solutions give us a good idea about the nature of convergence including the convergence time in a stochastic environment.

Since the extended least squares (ELS) and recursive least squares (RLS) algorithms can be treated as simplified and approximate versions of the RML algorithm [4], the analytical method presented here is applicable to these algorithms also. Although the hyperstable adaptive recursive filter (HARF) and its modified version (SHARF) [2, 16] are based on a different reasoning, recursions of these algorithms are nearly the same as those of the RLS algorithm but for using a filtered gradient. Hence the analytical procedure given here is applicable for these algorithms also.

EXPERIMENTAL RESULTS

To examine the validity of the present analysis, the present analytical procedure is applied to several adaptive filtering problems using different algorithm parameters and SNR's. Here we present some of the representative results obtained from the following example of the system to be identified. In this example the output of a second order IIR system driven by pseudo random gaussian noise is considered as the deterministic part of the noisy signal for prediction. The transfer function of the system considered here is

$$H(z) = \frac{1 - [\,2r_1\cos(0.2\,\pi)]\,z^{-1} + r_1^2 z^{-2}}{1 - [\,2r_2\cos(0.3\,\pi)]\,z^{-1} + r_2^2 z^{-2}}$$

where the normalized frequency and radius of the conjugate poles are 0.15 and r_2 respectively and the normalized frequency and radius of the conjugate zeros are 0.1 are r_1 respectively. Results are obtained for $r_1 = 0.8$ and $r_2 = 0.9$ unless specified, otherwise. Further, the variance of the input noise = 1.0. Independent and appropriate amount of pseudo random gaussian noise is added to this system output to form a noisy signal at different SNR's. Here for all the results an SNR of 1000.0 is considered.

Figures 2*a* and 2*b* illustrate the simulated and theoretical pole and zero parameter trajectories of the adaptive algorithm for $\mu = 0.02$ and $\lambda = 0.98$. Figures 3 and 4 are obtained for the values of μ of 0.05 and 0.1, keeping the value of λ at 0.98. Figures 3 and 4 illustrate only the pole parameters. Here, the simulation results represent the individual parameter trajectories only. The error covariance matrix, P_t is initialised with $\alpha = 0.01$. All these results show that there is good agreement between the theoretical and simulation results. Because of the flat nature of the criterion function, away from the optimum, during the initial convergence, the covergence speed is low. During the intermediate convergence region it is faster. Convergence speed during the final convergence region falls in between. It is also observed that convergence speed increases as μ increases but results in high parameter noise variance. From Figs 2*a*, 3 and 4 it is seen that parameter c_1 attains the near optimum value in around 3300, 1900 and 1500 iterations respectively.

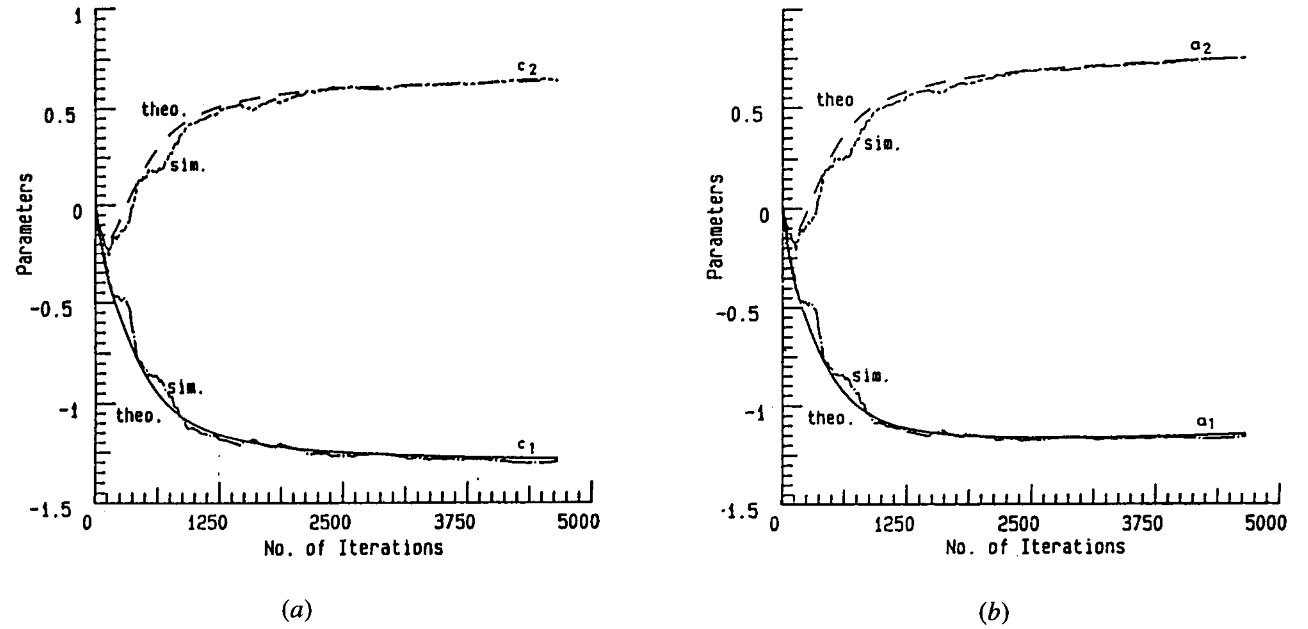

Fig 2 Theoretical and simulated (*a*) pole and (*b*) zero parameter trajectories (inverse model) of the adaptive algorithm for $\mu = 0.02$, $\lambda = 0.98$

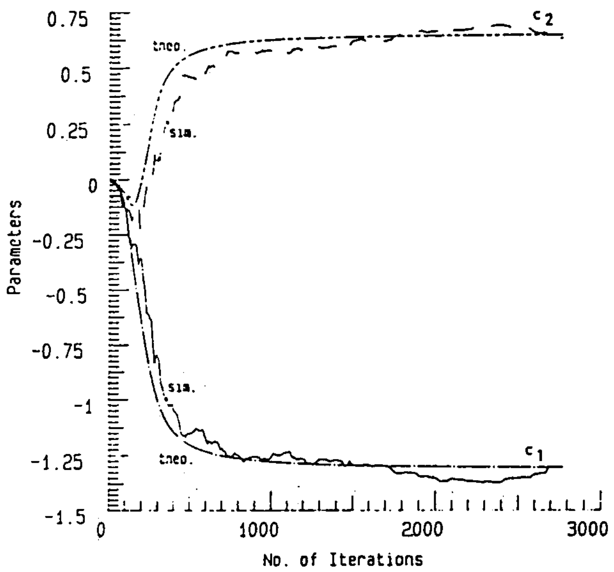

Fig 3 Pole parameter trajectories of the adaptive algorithm for $\mu = 0.05$, $\lambda = 0.98$

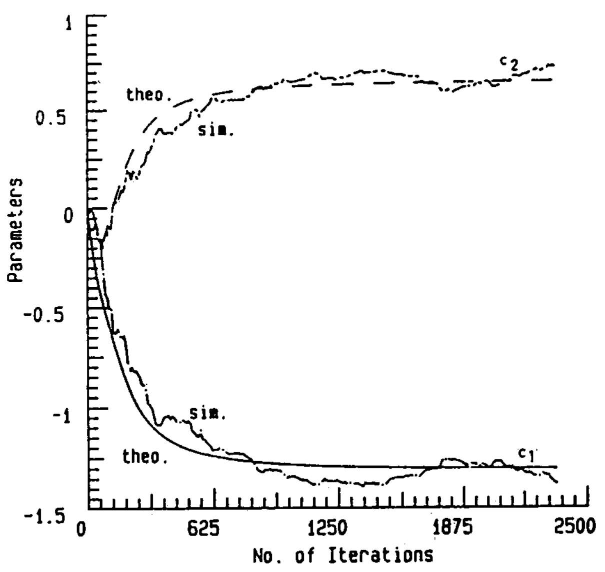

Fig 4 Pole parameter trajectories of the adaptive algorithm for $\mu = 0.1$, $\lambda = 0.98$

Figures 5 and 6 illustrate agreement the pole parameter trajectories for values of λ of 0.99 and 0.95 respectively, keeping the value of μ at 0.02. Here also good agreement is observed between theoretical and simulation results. It is also seen that convergence speed is proportional to $(1 - \lambda)$. From the Figs 5 and 6 it is seen that the parameter c_1 attains a near optimum value in around 5000 and 2500 iterations respectively. But increased $(1 - \lambda)$ causes more parameter noise variance. From all the results, it is seen that the applicability of the ODE approach is better for smaller values of μ and $(1 - \lambda)$.

The convergence speed of an identification algorithm for pole-zero systems largely depends on the nature of the performance surface. The non-quadratic nature of a performance function depends on how close the poles and zeros of the system to the unit circle. To study the dependence of the convergence nature on the pole radius, we considered different values for r_1 and r_2 in the system transfer function. Here, we include results obtained for $r_1 = r_2 = 0.9$, 0.95 and 0.98. Though results are obtained for all parameters, to avoid redundancy, only the result of a single parameter is shown in Fig 7. In general, the convergence time of a gradient algorithm increases with the closeness of the pole radius to the unit circle [10]. But, in the case of the present algorithm, it is seen that there is not much difference in convergence time for the different pole radii as illustrated in the figure. This is one of the advantages of the RML algorithm.

With this analysis, more general results can not be obtained like in a closed form analytical expression.This is mainly because, the convergence performance largely depend on the transfer function, its pole and zero locations and the input SNR. However, for a given class of systems some general reuslts can be obtained as shown here.

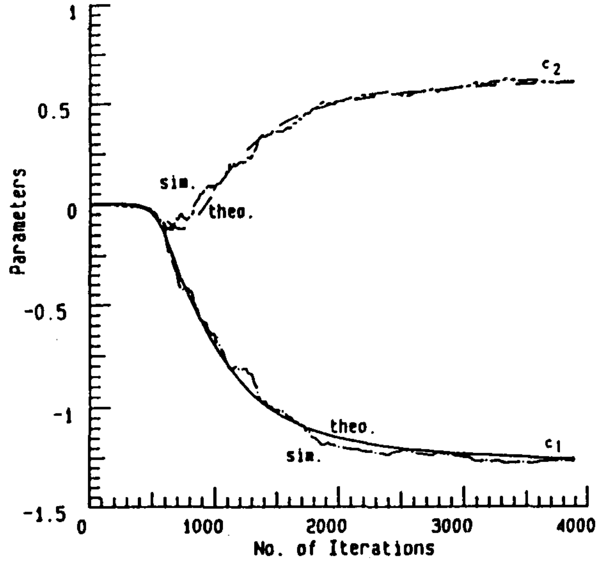

Fig 5 Pole parameter trajectories of the adaptive algorithm for $\mu = 0.02$, $\lambda = 0.99$

CONCLUSIONS

In this paper, the convergence behavior of adaptive algorithms for identification of pole-zero systems is analyzed. A method of solving the ODEs which represents the ensemble behavior of the adaptive identification algorithms is applied to make this study. Further, dependence of the convergence behavior on algorithm parameters (μ and λ), system pole radius are studied. This study is carried out with the help of both analytical and simulation results.

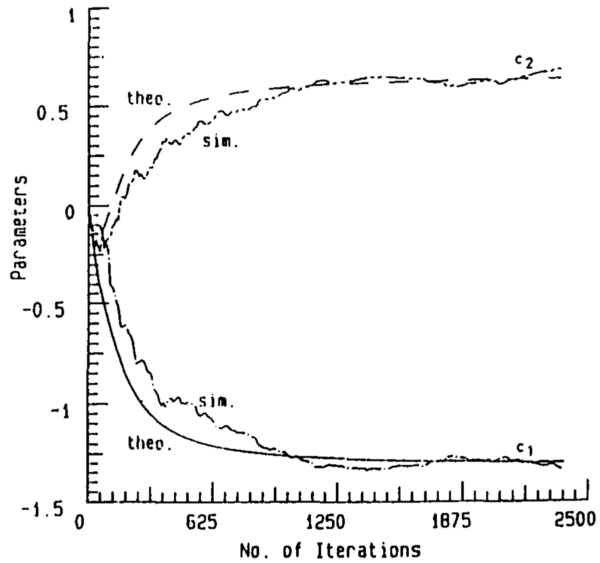

Fig 6 Pole parameter trajectories of the adaptive algorithm for $\mu = 0.02$, $\lambda = 0.95$

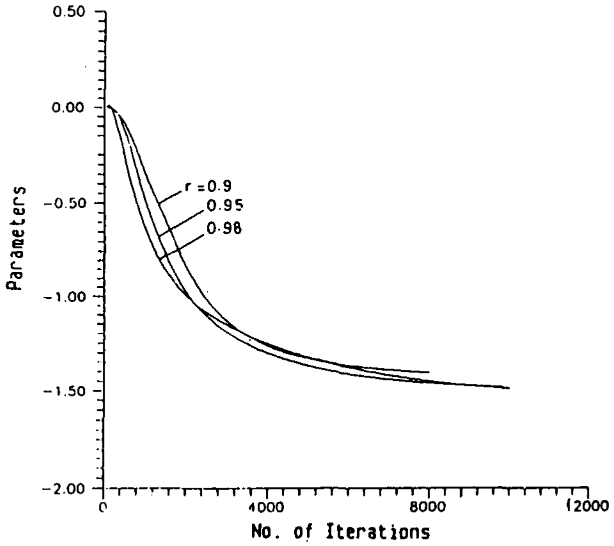

Fig 7 Pole parameter trajectories of the adaptive algorithm for different pole radii

In this paper, it is seen that there is a good agreement between the theoretical and simulation parameter trajectories. Because of the nonquadratic nature of the criterion function, it is found that the convergence speed is low during the initial convergence region, faster during the intermediate convergence region and it is in between during final convergence. Further, it is observed that the convergence speed is proportional to $\mu\ (1 - \lambda)$, but at the cost of parameter noise variance. It is also found that the applicability of the ODE approach is better for smaller values of $\mu(1 - \lambda)$.

The convergence speed of the adaptive algorithm for pole-zero system also depends on the nature of performance surface. Although the convergence time increases with the closeness of the pole radius to the unit circle, it is found that there is not much of difference in the convergence time for different values of radii.

REFERENCES

1. B Widrow & S D Stearns, *Adaptive Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1984.
2. C F N Cowan & P M Grant, *Adaptive Filters*, Prentice-Hall, Englewood Cliffs, NJ, 1985.
3. L Ljung & T Soderstrom, *Theory and Practice of Recursive Identification*, MIT Press, Cambridge, MA, 1983.
4. G C Goodwin & K S Sin, *Adaptive Filtering, Prediction and Control*, Prentice-Hall, Englewood Cliffs, NJ, 1984.
5. John J Shynk, Adaptive IIR filtering, *IEEE ASSP Magazine*, pp 4-21, April 1989.
6. C R Johnson, Jr, Adaptive IIR filtering: Current results and open issues, *IEEE Trans Inform Theory*, vol IT-30, no 2, pp 237-250, March 1984.
7. B Friedlander, System identification techniques for adaptive signal processing, *IEEE Trans Acoust, Speech and Singal Processing*, vol ASSP-30, pp 240-246, April 1982.
8. L Ljung, Analysis of recursive stochastic algorithms, *IEEE Trans Autom Contol*, vol AC-22, no 4, pp 551-575, August 1977.
9. R V Raja Kumar, Recursive Center-Frequency Adaptive Filters for the Enhancement and Tracking of Bandpass Signals, *PhD Thesis*, Department of E&ECE, IIT, Kharagpur, May 1986.
10. R V Raja Kumar & R N Pal, Performance analysis of the recursive center-frequency adaptive bandpass filters, *Signal Processing*, vol 17, no 2, pp 105-118, June 1989.
11. H Fan, Application of Benveniste's Convergence results in the study of adaptive IIR filtering algorithms, *IEEE Trans Inform Theory*, vol IT-34, no 4, pp 692-709, July 1988.
12. H J Kushner & A Shwartz, Weak convergence and asymptotic properties of adaptive filters with constant gains, *IEEE Trans Inform Theory*, vol IT-30, no 2, pp 177-182, March 1984.
13. A Benveniste, M Goursat & G Ruget, Analysis of stochastic approximation schemes with discontinuous and dependent forcing terms with applications to data communication algorithms, *IEEE Trans Automat control*, vol AC-25, no 6, pp 1042-1058, December 1980.
14. R V Raja Kumar & R N Pal, Tracking of bandpass signals using center-frequency adaptive filters, *IEEE Trans Acoust, Speech and Signal Processing*, vol ASSP-38, pp 1710-1721, October 1990.
15. D B Rao & R Peng, Tracking characteristics of the constrained IIR adaptive notch filter, *IEEE Trans Acoust, Speech and Signal Processing*, vol ASSP-36, pp 1466-1479, September 1988.
16. C R Johnson, Jr, M G Larimore, J R Treichler & B D O Anderson, SHARF Convergence Properties, *IEEE Trans Accoust, Speech and Signal Processing*, vol ASSP-29, pp 659-670, June 1981.
17. C B Rama Rao, Convergence and Tracking Behavioural Analysis of some IIR Adaptive Algorithms, *PhD Thesis*, Department of E&ECE, IIT Kharagpur, India, February 1994.
18. R V Raja Kumar & C B Rama Rao, Mean-convergence Behavioural Analysis of IIR Adaptive Filters, *Signal Processing Journal* (accepted), November 1994.