数理解析研究所講究録
第 562 巻 1985 年 187-207

適応格子形アルゴリズムの収束性の比較について

(Convergence Properties of Simplified Gradient Adaptive Lattice Algorithm)

京大工　飯国洋二　(Youji Iiguni)

1. Introduction

Using the gradient descendent technique, several gradient adaptive lattice (GAL) algorithms were developed [1] [2]. Compared with a non-lattice gradient least mean square (LMS) algorithm, these GAL algorithms generally converge faster, and their convergence rate is independent of the eigenvalue spread of the input data covariance matrix. Moreover, in digital computations, the roundoff noise of filter output due to parameter accurancies by quantization must be paid much attention. Several studies show that lattice filters can achieve a superior performance to a corresponding direct realization, on which the structure of the LMS algorithm is based. Thus, in this respect too, the GAL algorithms are expected to have good properties.

From the above reasons, the GAL algorithms have received much attention as a linear prediction algorithm. But in computational complexity the GAL algorithms are not so simple as the LMS algorithm. Therefore, to be free from hardware complexity and reduce an amount of computation time, several simplified versions of the GAL algorithm, what is called, the simplified GAL (SGAL) algorithms were proposed [3]. In these algorithms, PARCOR coefficients are computed using a sign-converter. These SGAL algorithms were applied to a vocoder and excellent results were obtained. Though the SGAL algorithms hold good features accompanied with the lattice structure, it doesn't match for the GAL algorithm in the convergence properties. Therefore in the SGAL algorithm it is rather important how we set a step size parameter.

Several results concerning with the convergence properties of the gradient type algorithms, i.e., the LMS algorithms including their simplified versions and the GAL algorithms, have been carried out so far [4] [5] [6] [7]. In this paper, assuming that the input signal is a stationary random process, convergence properties of the SGAL algorithms are analyzed. First, convergence models are derived, which predict the mean value trajectories of the estimated PARCOR coefficents and give us the convergence rate. Secondly, the variances of the estimation error of the PARCOR coefficients in steady states are calculated. Lastly, the performances of these algorithms, i.e., the convergence rates and the error variances, are compared. These results are available on the choice of the step size. Numerical simulations are performed to show the validity of our analysis.

2. Algorithms

Let x_t be a scalar stationary Gaussian process with zero mean and ρ^{p+1} $(p=0,1,\cdots)$ PARCOR coefficients which play an important pole in lattice algorithms.

Now we discuss the six recursive algorithms for estimation of ρ^{p+1}. A basic adaptive lattice structure is given by

$$e_t^0=r_t^0=x_t$$

$$e_t^{p+1}=e_t^p+\rho_t^{p+1}r_{t-1}^p \tag{2.1}$$

$$r_t^{p+1}=r_{t-1}^p+\rho_t^{p+1}e_t^p \quad , \tag{2.2}$$

where e_t^p and r_t^p are the forward and backward prediction errors for p-th order, respectively. The estimated PARCOR coefficient ρ_t^{p+1} is given by either the following equations:

A) $\rho_{t+1}^{p+1}=\rho_t^{p+1}-\frac{\alpha}{2}(e_t^{p+1}r_{t-1}^p+r_t^{p+1}e_t^p)$ (2.3)

B) $\rho_{t+1}^{p+1}=\rho_t^{p+1}-\alpha r_t^{p+1}e_t^p$ (2.4)

C) $\rho_{t+1}^{p+1}=\rho_t^{p+1}-\alpha r_t^{p+1}\mathrm{sgn}(e_t^p)$ (2.5)

D) $\gamma_{t+1}^{p+1}=(1-\alpha)\gamma_t^{p+1}-\alpha\,\mathrm{sgn}(e_t^p)\mathrm{sgn}(r_{t-1}^p)$

$\rho_{t+1}^{p+1}=\sin(\frac{\pi}{2}\gamma_{t+1}^{p+1})$ (2.6)

E) $\rho_{t+1}^{p+1}=\rho_t^{p+1}-\alpha\,\mathrm{sgn}(r_t^{p+1})\mathrm{sgn}(e_t^p)$ (2.7)

F) $\rho_{t+1}^{p+1}=\rho_t^{p+1}-\alpha\,\mathrm{sgn}(r_t^{p+1})e_t^p$ (2.8)

and

$$\mathrm{sgn}(x)=\begin{cases}+1 & x\geq 0\\ -1 & x<0\end{cases} \quad (2.9)$$

where α is a small positive constant, called the step size parameter. When we run these algorithms, we must adjust the step size adequately such that our desirable convergence properties can be obtained. The first and second algorithms are the GAL algorithms. The others are SGAL algorithms in which PARCOR coefficients are computed using a sign converter. The algorithm D seems to take much time to compute the sine function. But preparing the corresponding table of the sine function, the computational time can be fairly reduced.

We assume that the mean square value of e_t^p is equal to that of r_{t-1}^p and write it as $\sigma_{p,t}^2$, i.e.,

$$E[(e_t^p)^2]=E[(r_{t-1}^p)^2]=\sigma_{p,t}^2\ , \quad (2.10)$$

where E[•] denotes the expectation operator. Since ρ_t^{p+1} is slowly varying with respect to $e_t^p e_t^p$, the random variable $e_t^p e_t^p$ is almost independent of the random variables ρ_t^{p+1}. Therefore, taking the ensemble average of $e_t^p e_t^p \rho_t^{p+1}$ and applying the averaging principle [8], we approximately have

$$E[e_t^p e_t^p \rho_t^{p+1}]=E[e_t^p e_t^p]\cdot E[\rho_t^{p+1}] \ . \qquad (2.11)$$

Under these assumptions, all $E[\rho_t^{p+1}]$ in the above six algorithms converge towards

$$\rho^{p+1}=-E_\infty[e_t^p r_{t-1}^p]/E_\infty[e_t^p e_t^p], \qquad (2.12)$$

where $E_\infty[\cdot]$ denotes the expectation operator when $t\to\infty$.

3. Convergence Models

In this section, for each of the above six algorithms, we derive a convergence model to get the convergence rate. This model gives the ensemble average trajectories of the estimated PARCOR coefficients, which start from the initial values and converge towards their steady state values.

Let the forward and the backward prediction coefficients corresponding to the PARCOR coefficients $\{\rho_t^1, \rho_t^2, \cdots, \rho_t^{p+1}\}$ be $\{a_{1,t}^p, a_{2,t}^p, \cdots, a_{p,t}^p\}$ and $\{b_{1,t}^p, b_{2,t}^p, \cdots, b_{p,t}^p\}$, respectively. Then e_t^p and r_t^p can be expressed as

$$e_t^p=\sum_{i=0}^{p} a_{i,t}^p x_{t-i} \qquad (3.1)$$

$$r_t^p=\sum_{i=0}^{p} b_{i,t}^p x_{t-p+i} \ . \qquad (3.2)$$

Order updating recursions for $a^p_{i,t}$ and $b^p_{i,t}$ are given by

$$a^{p+1}_{i,t}=a^p_{i,t}+\rho^{p+1}_t b^p_{p+1-i,t-1} \qquad a^p_{0,t}=1 \qquad (3.3)$$

$$b^{p+1}_{i,t}=b^p_{i,t-1}+\rho^{p+1}_t a^p_{p+1-i,t} \qquad b^p_{0,t}=1 \ . \qquad (3.4)$$

From (3.1) and (3.2), we have

$$E[e^p_t e^p_t]=\sum_{i=0}^{p}\sum_{j=0}^{p} E[a^p_{i,t}]\ E[a^p_{j,t}]\ R_{i-j} \qquad (3.5)$$

$$E[e^p_t r^p_{t-1}]=\sum_{i=0}^{p}\sum_{j=0}^{p} E[a^p_{i,t}]\ E[b^p_{j,t-1}]\ R_{p+1-i-j} \qquad (3.6)$$

where $R_i=E[x_t x_{t+i}]$.

Since $(b^p_{p+1-i,t-1},\ a^p_{p+1-i,t})$ and ρ^{p+1}_t are uncorrelated, we have

$$E[a^{p+1}_{i,t}]=E[a^p_{i,t}]+E[\rho^{p+1}_t]E[b^p_{p+1-i,t-1}] \qquad E[a^p_{0,t}]=1 \qquad (3.7)$$

$$E[b^{p+1}_{i,t}]=E[b^p_{i,t-1}]+E[\rho^{p+1}_t]E[a^p_{p+1-i,t}] \qquad E[b^p_{0,t}]=1\ . \qquad (3.8)$$

Let k^{p+1}_t be defined as

$$k^{p+1}_t \triangleq -E[e^p_t r^p_{t-1}]/E[e^p_t e^p_t]. \qquad (3.9)$$

It is clear that k^{p+1}_t converges towards ρ^{p+1} when $t\to\infty$.

A): Taking an ensemble average of both sides of (2.3) and applying the averaging principle to it, we have [4]

$$E[\rho^{p+1}_{t+1}]=(1-\alpha\sigma^2_{p,t})E[\rho^{p+1}_t]+\alpha k^{p+1}_t\sigma^2_{p,t}. \qquad (3.10)$$

If $E[\rho^{p+1}_t]$ are given, k^{p+1}_t and $\sigma^2_{p,t}$ can be computed from (3.5)-(3.9), and then $E[\rho^{p+1}_{t+1}]$ from (3.10). Using these manipulations iteratively, we can get the mean value trajectories of the estimated PARCOR coefficients, i.e., the convergence model.

B): Similarly we have the same convergence model as (3.10).

C): We assume that e_t^p and r_{t-1}^p are jointly Gaussian with zero mean. Defining the joint distribution density function of e_t^p and r_{t-1}^p as $f(e_t^p, r_{t-1}^p)$, it can be expressed as

$$f(e_t^p, r_{t-1}^p) = \frac{1}{2\pi\sigma_{p,t}^2\sqrt{1-(k_t^{p+1})^2}} \exp\left\{-\frac{(e_t^p)^2+(r_{t-1}^p)^2+2k_t^{p+1}e_t^p r_{t-1}^p}{2\sigma_{p,t}^2(1-(k_t^{p+1})^2)}\right\}. \qquad (3.11)$$

Using (3.11), we have

$$E[r_t^{p+1}\mathrm{sgn}(e_t^p)]=E[r_{t-1}^p\mathrm{sgn}(e_t^p)]+E[\rho_t^{p+1}]E[e_t^p\mathrm{sgn}(e_t^p)]$$

$$= \int_0^\infty de_t^p \int_{-\infty}^\infty (r_{t-1}^p+E[\rho_t^{p+1}]e_t^p)\ f(e_t^p, r_{t-1}^p)\ dr_{t-1}^p$$

$$- \int_0^\infty de_t^p \int_{-\infty}^\infty (r_{t-1}^p-E[\rho_t^{p+1}]e_t^p)\ f(-e_t^p, r_{t-1}^p)\ dr_{t-1}^p$$

$$=\sqrt{\frac{2}{\pi}}\,(E[\rho_t^{p+1}]-k_t^{p+1})\sigma_{p,t}\ . \qquad (3.12)$$

Hence the convergence model becomes

$$E[\rho_{t+1}^{p+1}]=(1-\alpha\sqrt{\frac{2}{\pi}}\sigma_{p,t})E[\rho_t^{p+1}]+\alpha\sqrt{\frac{2}{\pi}}\,k_t^{p+1}\sigma_{p,t}. \qquad (3.13)$$

D): Similarly we have

$$E[\mathrm{sgn}(e_t^p)\mathrm{sgn}(r_{t-1}^p)]$$

$$=2\int_0^\infty\int_0^\infty f(e_t^p, r_{t-1}^p)\ de_t^p\ dr_{t-1}^p\ -2\int_0^\infty\int_0^\infty f(-e_t^p, r_{p-1}^p)\ de_t^p\ dr_{t-1}^p$$

$$=-\frac{2}{\pi}\,\mathrm{Sin}^{-1}(k_t^{p+1}) \qquad (3.14)$$

so that

$$E[\gamma_{t+1}^{p+1}]=(1-\alpha)E[\gamma_t^{p+1}]+\alpha\frac{2}{\pi}\,\mathrm{Sin}^{-1}(k_t^{p+1})\ . \qquad (3.15)$$

Noting that

$$E[\sin(\frac{\pi}{2}\gamma_t^{p+1})-\sin(\frac{\pi}{2}E[\gamma_t^{p+1}])]$$

$$=2E[\cos\frac{\pi}{4}(\gamma_t^{p+1}+E[\gamma_t^{p+1}])\cdot\sin\frac{\pi}{4}(\gamma_t^{p+1}-E[\gamma_t^{p+1}])]$$

$$\cong 2\cos\frac{\pi}{2}E[\gamma_t^{p+1}]\cdot E[\gamma_t^{p+1}-E[\gamma_t^{p+1}]]=0 \ , \qquad (3.16)$$

we have

$$E[\rho_{t+1}^{p+1}]=\sin(\frac{\pi}{2}E[\gamma_{t+1}^{p+1}]). \qquad (3.17)$$

The equation (3.15) and (3.17) give the convergence model of the algorithm D.

E): Also in this algorithm, the convergence model can be obtained in the same way as stated in the above. But here we present a rather detailed derivation for later use. We define the joint distribution density function of (e_t^p, r_{t-1}^p, ρ_t^{p+1}) and the distribution density function of ${}_t^{p+1}$ as g(e_t^p, r_{t-1}^p, ρ_t^{p+1}) and h(ρ_t^{p+1}), respectively. Since (e_t^p, r_{t-1}^p) and ρ_t^{p+1} are almost uncorrelated, we can have $g\cong f\cdot h$. Let G(e_t^p, r_{t-1}^p, ρ_t^{p+1}) denote the random variable expressed as the function of e_t^p, r_{t-1}^p and ρ_t^{p+1}. Then the expectation of G becomes

$$E[G(e_t^p,\ r_{t-1}^p,\ \rho_t^{p+1})]=\iiint G\cdot g\ de_t^p\ dr_{t-1}^p\ d\rho_t^{p+1}$$

$$\cong\int\left[\iint G\cdot f\ de_t^p\ dr_{t-1}^p\right]h\ d\rho_t^{p+1}$$

$$=E[\iint G\cdot f\ de_t^p\ \ dr_{t-1}^p\ ]_{\rho_t^{p+1}} \ , \qquad (3.18)$$

where $E[\cdot]_{\rho_t^{p+1}}$ is the expectation operator of ρ_t^{p+1}. Using this property, the probability of $\mathrm{sgn}(r_t^{p+1})\mathrm{sgn}(e_t^p)=1$ becomes

$$E\left[\iint_{(r_{t-1}^p+\rho_t^{p+1}e_t^p)e_t^p>0} f(e_t^p,\ r_{t-1}^p)\ de_t^p\ dr_{t-1}^p\right]_{\rho_t^{p+1}} \ , \qquad (3.19)$$

After some manipulations, the integral (3.19) reduces to

$$\frac{1}{2}+\frac{1}{\pi}E\left[\mathrm{Tan}^{-1}\left(\frac{\rho_t^{p+1}-k_t^{p+1}}{\sqrt{1-(k_t^{p+1})^2}}\right)\right]\rho_t^{p+1}. \tag{3.20}$$

We can establish the following approximation:

$$E\left[\mathrm{Tan}^{-1}\left(\frac{\rho_t^{p+1}-k_t^{p+1}}{\sqrt{1-(k_t^{p+1})^2}}\right)-\mathrm{Tan}^{-1}\left(\frac{E[\rho_t^{p+1}]-k_t^{p+1}}{\sqrt{1-(k_t^{p+1})^2}}\right)\right]\rho_t^{p+1}$$

$$=E\left[\mathrm{Tan}^{-1}\left(\frac{\dfrac{\rho_t^{p+1}-E[\rho_t^{p+1}]}{\sqrt{1-(k_t^{p+1})^2}}}{1+\dfrac{(E[\rho_t^{p+1}]-k_t^{p+1})(\rho_t^{p+1}-k_t^{p+1})}{1-(k_t^{p+1})^2}}\right)\right]\rho_t^{p+1}$$

$$\cong\frac{E[\rho_t^{p+1}-E[\rho_t^{p+1}]]}{\sqrt{1-(k_t^{p+1})^2}\left(1+\dfrac{(E[\rho_t^{p+1}]-k_t^{p+1})^2}{1-(k_t^{p+1})^2}\right)}=0. \tag{3.21}$$

From this, (3.20) approximately becomes

$$\frac{1}{2}+\frac{1}{\pi}\mathrm{Tan}^{-1}\left(\frac{E[\rho_t^{p+1}]-k_t^{p+1}}{\sqrt{1-(k_t^{p+1})^2}}\right). \tag{3.22}$$

On the other hand, the probability of $\mathrm{sgn}(r_t^{p+1})\mathrm{sgn}(e_t^p)=-1$ can be similarly given by

$$\frac{1}{2}-\frac{1}{\pi}\mathrm{Tan}^{-1}\left(\frac{E[\rho_t^{p+1}]-k_t^{p+1}}{\sqrt{1-(k_t^{p+1})^2}}\right). \tag{3.23}$$

From the above, we have

$$E[\rho_{t+1}^{p+1}]=E[\rho_t^{p+1}]-\alpha E[\mathrm{sgn}(r_t^{p+1})\mathrm{sgn}(e_t^p)]$$

$$=E[\rho_t^{p+1}]+\frac{2\alpha}{\pi}\mathrm{Tan}^{-1}\left(\frac{k_t^{p+1}-E[\rho_t^{p+1}]}{\sqrt{1-(k_t^{p+1})^2}}\right) \quad . \qquad (3.24)$$

F): Using the same technique as in (3.21), we have

$$E[\mathrm{sgn}(r_t^{p+1})e_t^p]$$

$$=2E\left[\int_{-\infty}^{\infty}dr_{t-1}^p \int_{r_{t-1}^p+\rho_t^{p+1}e_t^p>0} e_t^p\cdot f(e_t^p,\ r_{t-1}^p)\ de_t^p\right]\rho_t^{p+1}$$

$$=\sqrt{\frac{2}{\pi}}\ \frac{\alpha\sigma_{p,t}\ (E[\rho_t^{p+1}]-k_t^{p+1})}{\sqrt{1-2E[\rho_t^{p+1}]k_t^{p+1}+E^2[\rho_t^{p+1}]}} \quad , \qquad (3.25)$$

so that

$$E[\rho_{t+1}^{p+1}]=\left(1-\sqrt{\frac{2}{\pi}}\frac{\alpha\sigma_{p,t}}{\sqrt{1-2E[\rho_t^{p+1}]k^{p+1}+E^2[\rho_t^{p+1}]}}\right)E[\rho_t^{p+1}]$$

$$+\sqrt{\frac{2}{\pi}}\cdot\frac{\alpha\sigma_{p,t}}{\sqrt{1-2E[\rho_t^{p+1}]k^{p+1}+E^2[\rho_t^{p+1}]}}\ k_t^{p+1} \quad . \qquad (3.26)$$

4. Variance of the PARCOR coefficients error

The estimated PARCOR coefficients ρ_t^{p+1} computed by the algorithms A-F fluctuate around ρ^{p+1} in a steady state. In this section, we derive the variance of this fluctuation for each algorithm.

A): Let rewrite (2.3) as

$$\rho_{t+1}^{p+1}-\rho^{p+1}=(1-\frac{\alpha}{2}(e_t^pe_t^p+r_{t-1}^pr_{t-1}^p))(\rho_t^{p+1}-\rho^{p+1})$$

$$-\alpha(e_t^pr_{t-1}^p+\frac{1}{2}\rho^{p+1}(e_t^pe_t^p+r_{t-1}^pr_{t-1}^p)). \qquad (4.1)$$

Squaring both sides and taking the ensemble average of them, the cross term of the right side becomes zero by the averaging principle. Hence the variance of the estimation error of the PARCOR coefficient, $Var_\infty(\rho_t^{p+1})$, can be expressed as follows [5]:

$$Var_\infty(\rho_t^{p+1}) = \lim_{t\to\infty} E[(\rho_t^{p+1}-\rho^{p+1})^2]$$

$$= \frac{(1-(\rho^{p+1})^2)^2\sigma_p^2\alpha}{2-\sigma_p^2\alpha(2+(\rho^{p+1})^2)} \cong \frac{\alpha\sigma_p^2}{2}(1-(\rho^{p+1})^2)^2 \qquad (4.2)$$

B and C): As for B and C, in the same way as in A, we have

$$Var_\infty(\rho_t^{p+1}) = \frac{(1-(\rho^{p+1})^2)\sigma_p^2\alpha}{2-3\sigma_p^2\cdot\alpha} \cong \frac{\alpha\sigma_p^2}{2}(1-(\rho^{p+1})^2) \qquad (4.3)$$

and

$$Var_\infty(\rho_t^{p+1}) = \frac{(1-(\rho^{p+1})^2)\sigma_p^2\alpha}{2\sqrt{\frac{2}{\pi}}+\sigma_p\cdot\alpha} \cong \sqrt{\frac{\pi}{8}}\ \sigma_p^2(1-(\rho^{p+1})^2)\alpha \qquad (4.4)$$

respectively.

D): Upon setting $\gamma^p = \frac{2}{\pi}\sin^{-1}(\rho^p)$, we can easily get the error variance of γ_t^{p+1} as follows:

$$Var_\infty(\gamma_t^{p+1}) = \frac{\alpha}{2-\alpha}(1-(\gamma^{p+1})^2) \cong \frac{\alpha}{2}(1-(\gamma^{p+1})^2) \quad . \qquad (4.5)$$

Since the following approximation can be established in a steady state

$$\rho_t^{p+1}-\rho^{p+1} \cong \frac{\pi}{2}(\gamma_t^{p+1}-\gamma^{p+1})\cos(\frac{\pi}{2}\gamma^{p+1}) \ , \qquad (4.6)$$

we have

$$Var_\infty(\rho_t^{p+1}) \cong \frac{\pi^2}{4} Var_\infty(\gamma_t^{p+1})(1-(\rho^{p+1})^2) \quad . \qquad (4.7)$$

Substituting (4.5) into (4.7) yields

$$\mathrm{Var}_{\infty}(\rho_t^{p+1}) = \frac{\pi^2}{8}\alpha(1-(\rho^{p+1})^2)(1-(\gamma^{p+1})^2) \quad . \qquad (4.8)$$

E): Since ρ_t^{p+1} is included in an argument of a sign function in (2.7), we can not explicitly get ρ_t^{p+1} out of the sign function. Therefore we can't apply the same technique as stated in the above. Since, in this case, $\rho_t^{p+1}-\rho_{t-1}^{p+1}=\alpha$ or $-\alpha$, we regard the behavior of ρ_t^{p+1} in the steady state as a discrete Markov process and write a probability of $\rho_t^{p+1}=i\alpha$ as P_i, where i is an integer and we omit the superscript p for simplicity. If $\rho_{t+1}^{p+1}=i\alpha$, $\mathrm{sgn}(r_t^{p+1})\mathrm{sgn}(e_t^p)=1$ when $\rho_t^{p+1}=(i+1)\alpha$ or $\mathrm{sgn}(r_t^{p+1})\mathrm{sgn}(e_t^p)=-1$ when $\rho_t^{p+1}=(i-1)\alpha$. Then from (3.19) and (3.20) we can establish the following recurrence formula in the steady state.

$$P_i = \left[\frac{1}{2} + \frac{1}{\pi}\mathrm{Tan}^{-1}\left(\frac{(i+1)\alpha-\rho^{p+1}}{\sqrt{1-(\rho^{p+1})^2}}\right)\right]P_{i+1} + \left[\frac{1}{2} - \frac{1}{\pi}\mathrm{Tan}^{-1}\left(\frac{(i-1)\alpha-\rho^{p+1}}{\sqrt{1-(\rho^{p+1})^2}}\right)\right]P_{i-1} \quad . \qquad (4.9)$$

Note that

$$\left|\frac{i\alpha-\rho^{p+1}}{\sqrt{1-(\rho^{p+1})^2}}\right| \ll \frac{\pi}{2} \qquad (4.10)$$

for all most i in the steady state. Upon applying $\mathrm{Tan}^{-1}(x)\cong x$ to (4.9), it reduces to

$$P_i = \left(\frac{1}{2} + \frac{(i+1)\alpha-\rho^{p+1}}{\sqrt{1-(\rho^{p+1})^2}}\right)P_{i+1} + \left(\frac{1}{2} - \frac{(i-1)\alpha-\rho^{p+1}}{\sqrt{1-(\rho^{p+1})^2}}\right)P_{i-1}, \qquad (4.11)$$

Then the solution of (4.11) is given as follows:

$$P_i = \frac{A}{\Gamma(a-ac+i)\cdot\Gamma(ac-i)} , \tag{4.12}$$

where

$a=\pi\sqrt{1-(\rho^{p+1})^2}/\alpha$

$c=0.5+\rho^{p+1}/a\cdot\alpha$

A: a normalized constant

Γ: Gamma function such that $\Gamma(x)=\int_0^\infty e^{-t}t^{x-1}dt$.

As long as (4.10) holds, it is guaranteed that $a-ac+i \gg 0$ and $ac-i \gg 0$. Using the approximation that $\Gamma(x+1)\cong x^x e^{-x}\sqrt{2\pi x}$, (4.12) including the normalized constant reduces to

$$P_i = \frac{1}{\sqrt{\pi(a-2)/2}} \exp\left[-\frac{(i\alpha-\rho^{p+1})^2}{(a-2)\alpha^2/2}\right] . \tag{4.13}$$

This equation shows that ρ_t^{p+1} is Gaussian with mean ρ^{p+1} and variance $(a-2)\alpha^2/4$. Therefore we have

$$\begin{aligned} Var_\infty(\rho_t^{p+1}) &= (a-2)\alpha^2/4 = \frac{\sqrt{1-(\rho^{p+1})^2}}{4}\alpha - \frac{\alpha^2}{8} \\ &\cong \frac{\sqrt{1-(\rho^{p+1})^2}}{4}\alpha . \end{aligned} \tag{4.14}$$

F): Also in this algorithm ρ_t^{p+1} is included in an argument of a sign function. But by quantizing e_t^p in step size Δ, we can treat this algorithm in the same way as in E. Upon setting $e_t^p=j\Delta$ and $\rho_t^{p+1}=i\Delta\alpha$, we approximately have

$$\Gamma(j,i)=Pr\{r_{t-1}^p+(i\Delta\alpha)e_t^p>0,\ e_t^p=j\Delta\}$$

$$= \int_{j\Delta}^{(j+1)\Delta} de_t^p \int_{r_{t-1}^p+(i\Delta\alpha)e_t^p>0} f(e_t^p,\ r_{t-1}^p)\ dr_{t-1}^p$$

$$\cong \frac{\Delta}{2\pi\sigma_p} \exp\left[-\frac{(j\Delta)^2}{2\sigma_p^2}\right]\left[\frac{2(\rho^{p+1}-i\Delta\alpha)(j\Delta)}{\sigma_p\sqrt{1-(\rho^{p+1})^2}}+\sqrt{2\pi}\right] . \qquad (4.15)$$

Writing the probability of $\rho_t^{p+1}=i\Delta\alpha$ as P_i, we have

$$P_i=\sum_{j=-\infty}^{\infty} \Gamma(j+i,-j)P_{j+i} . \qquad (4.16)$$

It is very difficult to derive the solution of this recurrence formula. Therefore we assume that ρ_t^{p+1} is Gaussian with mean ρ^{p+1} and variance d^2, i.e.,

$$P_i=A \exp\left[-\frac{(i\Delta\alpha-\rho^{p+1})^2}{2d^2}\right] . \qquad (4.17)$$

Substituting (4.17) into (4.16) gives

$$\mathrm{Var}_\infty(\rho_t^{p+1})=d^2=\sqrt{\frac{\pi}{8}(1-(\rho^{p+1})^2)}\,\sigma_p\alpha . \qquad (4.18)$$

The derivation is presented in Appendix.

5. Simulations

Numerical simulations were performed to support the validity of our methods described in sections 3 and 4. Here we mention only about the algorithm E. Using the data generated by the eighth autoregressive process whose PARCOR coefficients correspond to $(\rho^1, \rho^2, \cdots, \rho^8)=(0.9, 0.8, \cdots, 0.2)$, we estimated PARCOR coefficents by the algorithm E with a step size $\alpha=1/32$.

Fig.1 depicts the convergence models, i.e., the ensemble average trajectories of $E[\rho_t^{p+1}]$, for p=1, 2, 4, 8. Curve 1 is the trajectory of our proposed convergence model (3.21) and Curve 2 is the simulated trajectory obtained by averaging 100 realizations. We find that our model predicts very closely the experimental results. The bias between Curve 1 and 2 for the first order is due to the approximation (2.10). In higher order cases, the bias arises mainly due to the fact that the fluctuation of the PARCOR coefficients in the previous stages have changed the stochastic properties of the following stage inputs.

On the other hand, Fig. 2 presents the error variance $Var_{\infty}(\rho_t^{p+1})$ of the simulated and theoretical values. The black line shows the theoretical values computed by (4.14) and the numbers enclosed with a circle indicate the plots given by the simulation results. It can be seen that our theoretical results agree with simulation results, where the number corresponds to the order of PARCOR coefficient. To get the simulated results, we also carried out the time averaging on the results from t=501 to t=1500 when t is sufficiently large such that ρ_t^{p+1} has converged.

6. Performance comparison

The convergence properties of six adaptive lattice algorithms, i.e., convergence models and parameter error variances in steady states are investigated so far. In this section, we compare the performance of these algorithms.

Assuming that $|E[\rho_t^{p+1}]-k_t^{p+1}|\ll 1$, we apply several rough approximations to our convergence models. Then they can be reduced to the following simpler form:

$$E[\rho_{t+1}^{p+1}]=(1-\mu)E[\rho_t^{p+1}]+\mu k_t^{p+1} \quad . \tag{6.1}$$

For example, as for algorithm F, we can approximate (3.23) as

$$E[\mathrm{sgn}(r_t^{p+1})e_t^p] = \sqrt{\frac{2}{\pi}} \frac{\alpha\sigma_{p,t}}{\sqrt{1-(k_t^{p+1})^2}} \frac{E[\rho_t^{p+1}]-k_t^{p+1}}{\sqrt{1+\frac{(E[\rho_t^{p+1}]-k_t^{p+1})^2}{1-(k_t^{p+1})^2}}}$$

$$\cong \sqrt{\frac{2}{\pi}} \frac{\alpha\sigma_{p,t}}{\sqrt{1-(k_t^{p+1})^2}} (E[\rho_t^{p+1}]-k_t^{p+1}) \ . \qquad (6.2)$$

Hence μ can be simply expressed as

$$\mu = \sqrt{\frac{2}{\pi}} \frac{\alpha\sigma_{p,t}}{\sqrt{1-(k_t^{p+1})^2}} \ . \qquad (6.3)$$

The parameter μ determines the convergence rate. The larger μ becomes, the faster $E[\rho_t^{p+1}]$ converges towards ρ^{p+1}. The form of μ is summarized in Table 1 for six algorithms, where we omit the superscript p for simplicity. This table indicates that the convergence rate depends on not only the step size α but also input power σ_p^2 in the algorithms A,B,D,F, and PARCOR coefficient ρ^{p+1} in E,F.

The variance of the estimation error of the PARCOR coefficient, $\mathrm{Var}_\infty(\rho_t^{p+1})$, is also summarized in Table 1. It is easily found that $\mathrm{Var}_\infty(\rho_t^{p+1})$ becomes large in proportion to the magnitude of the step size. Therefore we must trade off between the convergence rate and the error variance when we choose the step size. The variance of the estimation error is affected by the input power and input statistics in almost all algorithms. It is reasonable that the algorithm C and F have the same convergence properties and the error variance though they are derived using the different techniques.

Now we choose the step size α so that all of these algorithms have the same convergence rate μ. These μ-normalized variance of PARCOR coefficients error $\overline{\mathrm{Var}}_{\infty}(\rho_t^{p+1})$ are shown in Table 1. Comparing $\overline{\mathrm{Var}}_{\infty}(\rho_t^{p+1})$ in magnitude, the performance of the algorithms can be obtained. For example, as for B and E with $\mu=1/32$ and $\rho^{p+1}=0.9$, $\overline{\mathrm{Var}}_{\infty}(\rho_t^{p+1})$ becomes $2.97\cdot10^{-3}$ and $7.32\cdot10^{-3}$, respectively. So it can be concluded that the performance of the algorithm B is about 2.5 times superior to that of E at the cost of increased computation time. It is also interesting to note that the performance of the algorithm E is slightly inferior to that of the algorithm C or F despite its simpler structure.

7. Conclusion

For each of four SGAL and two GAL algorithms, a convergence model that gives us a convergence rate, and a variance of the estimation error of the PARCOR coefficient are derived. The convergence rate and error variance are expressed as the functions of a step size α, a input power σ_p^2 and PARCOR coefficient ρ^{p+1}. Our analyses show how to determine a step size such that we can get a desirable performance especially in the SGAL algorithms.

References

[1] L. J. Griffiths, "A continuously adaptive filter implemented as a lattice structure," in Proc. IEEE Conf., Acoust., Speech, Signal Process., pp.683-686, May 1977.

[2] L. J. Griffiths, "An adaptive lattice structure for noise cancelling applications," in Proc. IEEE Conf., Acoust., Speech, Signal Process., pp.87-90, Apr. 1978.

[3] F. Sugiyama and M. Nakamura, "An LPC vocoder for efficient implementation," in Proc. IEEE Conf., Acoust., Speech, Signal Process., pp.355-358, Apr. 1981.

[4] M. L. Honig and D. G. Messerschmitt, "Convergence Properties of an Aadaptive Digital Lattice Filter," IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-29, no.3, June 1981.

[5] G.R.L. Sohie and L.H. Sibul, "Stochastic Convergence Properties of the Aadaptive Gradient Lattice," IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-32, no.1, Feb. 1984.

[6] D.J. Robertson and D.D. Falconer, "Analysis and Simulation of Improved Echo Canceller Adaptation Algorithms with Single-Bit Correlation Multipliers," IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-32, no.5, Oct. 1984.

[7] H. Sari, "Performance evaluation of three adaptive equalization algorithms," in Proc. IEEE Conf., Acoust., Speech, Signal Process., pp.1385-1389, 1982.

[8] C. Samson and V.U. Reddy, "Fixed Point Error Analysis of the Normalized Ladder Algorithm," IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-31, no.5, pp.1177-1191, Oct. 1983.

Appendix

Substituting (4.17) into (4.16) yields

$$P_i = \frac{1}{\sqrt{2\pi}\,\sigma_p}\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2\sigma_p^2} - \frac{1}{2d^2}(x\alpha + i\alpha\Delta - \rho^{p+1})^2\right]dx$$

$$+\frac{(i\Delta\alpha - \rho^{p+1})}{\pi\sigma_p^2\sqrt{1-(\rho^{p+1})^2}}\int_{-\infty}^{\infty}x\cdot\exp\left[-\frac{x^2}{2\sigma_p^2} - \frac{1}{2d^2}(x\alpha + i\Delta\alpha - \rho^{p+1})^2\right]dx$$

$$+\frac{\alpha}{\pi\sigma_p^2\sqrt{1-(\rho^{p+1})^2}}\int_{-\infty}^{\infty}x^2\cdot\exp\left[-\frac{x^2}{2\sigma_p^2} - \frac{1}{2d^2}(x\alpha + i\Delta\alpha - \rho^{p+1})^2\right]dx \,. \tag{A.1}$$

For simplicity, we define a, b, c as follows:

$$a^2 = 1/(2\sigma_p^2), \quad b^2 = \alpha^2/(2d^2), \quad c = (\rho^{p+1} - i\Delta\alpha)/\alpha \quad .$$

Then (A.1) leads to

$$\exp\left[-\frac{(i\Delta\alpha - \rho^{p+1})^2}{2d^2}\right] = \frac{1}{\sqrt{2\pi}\,\sigma_p}\sqrt{\frac{\pi}{a^2+b^2}}\exp\left(-\frac{a^2b^2c^2}{a^2+b^2}\right)$$

$$+\frac{(i\Delta\alpha - \rho^{p+1})\,b^2c}{\pi\sigma_p^2\sqrt{1-(\rho^{p+1})^2}}\cdot\sqrt{\frac{\pi}{(a^2+b^2)^3}}\exp\left(-\frac{a^2b^2c^2}{a^2+b^2}\right)$$

$$+\frac{\alpha}{\pi\sigma_p^2\sqrt{1-(\rho^{p+1})^2}}\left[\frac{1}{2}\sqrt{\frac{\pi}{(a^2+b^2)^3}} + \left(\frac{b^2c}{a^2+b^2}\right)^2\sqrt{\frac{\pi}{a^2+b^2}}\exp\left(-\frac{a^2b^2c^2}{a^2+b^2}\right)\right] \tag{A.2}$$

Comparing (2.8) with (2.5), the variance of the estimation error of the PARCOR coefficients computed by (2.8) is expected to be the same order in magnitude as the one by (2.5). Therefore it can be seen from (4.4) that

$$(\alpha\sigma_p/d)^2 \ll 1 \,. \tag{A.3}$$

Then we can approximately have

$$\exp\left(-\frac{a^2b^2c^2}{a^2+b^2}\right)$$

$$\cong \exp\left[-\frac{(i\Delta\alpha-\rho^{p+1})^2}{2d^2}\right]\left[1+\frac{\alpha^2\sigma_p^2}{2d^4}(i\Delta\alpha-\rho^{p+1})^2\right]. \qquad \text{(A.4)}$$

Using this approximation, (A.2) becomes

$$1-\frac{1}{\sqrt{2\pi}\,\sigma_p}\sqrt{\frac{\pi}{a^2+b^2}}-\frac{\alpha}{2\pi\sigma_p^2\sqrt{1-(\rho^{p+1})^2}}\sqrt{\frac{\pi}{(a^2+b^2)^3}}$$

$$+\left[\frac{1}{\sqrt{2\pi}}\sqrt{\frac{\pi}{a^2+b^2}}\;\frac{\alpha^2\sigma_p}{2d^2}-\frac{\alpha}{2\pi\sigma_p^2\sqrt{1-(\rho^{p+1})^2}}\sqrt{\frac{\pi}{(a^2+b^2)^3}}\right.$$

$$+\frac{\alpha^3}{4\pi d^2\sqrt{1-(\rho^{p+1})^2}}\sqrt{\frac{\pi}{(a^2+b^2)^3}}$$

$$\left.+\frac{\alpha^3}{4\pi\sigma_p^2 d^2\sqrt{1-(\rho^{p+1})^2}}\sqrt{\frac{\pi}{(a^2+b^2)^5}}\right]\cdot(i\Delta\alpha-\rho^{p+1})^2=0 \qquad \text{(A.5)}$$

Considering that (A.5) must be established for all i, we have

$$1=\frac{1}{\sqrt{2\pi}\,\sigma_p}\sqrt{\frac{\pi}{a^2+b^2}}+\frac{\alpha}{2\pi\sigma_p^2\sqrt{1-(\rho^{p+1})^2}}\sqrt{\frac{\pi}{(a^2+b^2)^3}} \qquad \text{(A.6)}$$

$$\frac{\alpha^2\sigma_p^2}{2d^2}=\frac{\alpha}{2\pi\sigma_p^2\sqrt{1-(\rho^{p+1})^2}}\sqrt{\frac{\pi}{(a^2+b^2)^3}}+\frac{\alpha^3}{4\pi\sigma_p^2 d^2\sqrt{(1-(\rho^{p+1})^2)}}\sqrt{\frac{\pi}{(a^2+b^2)^5}}\,. \qquad \text{(A.7)}$$

The second terms of the right hand side in both (A.6) and (A.7) can be neglected. Then, upon solving simultaneous linear equations for d^2, we can have (4.18).

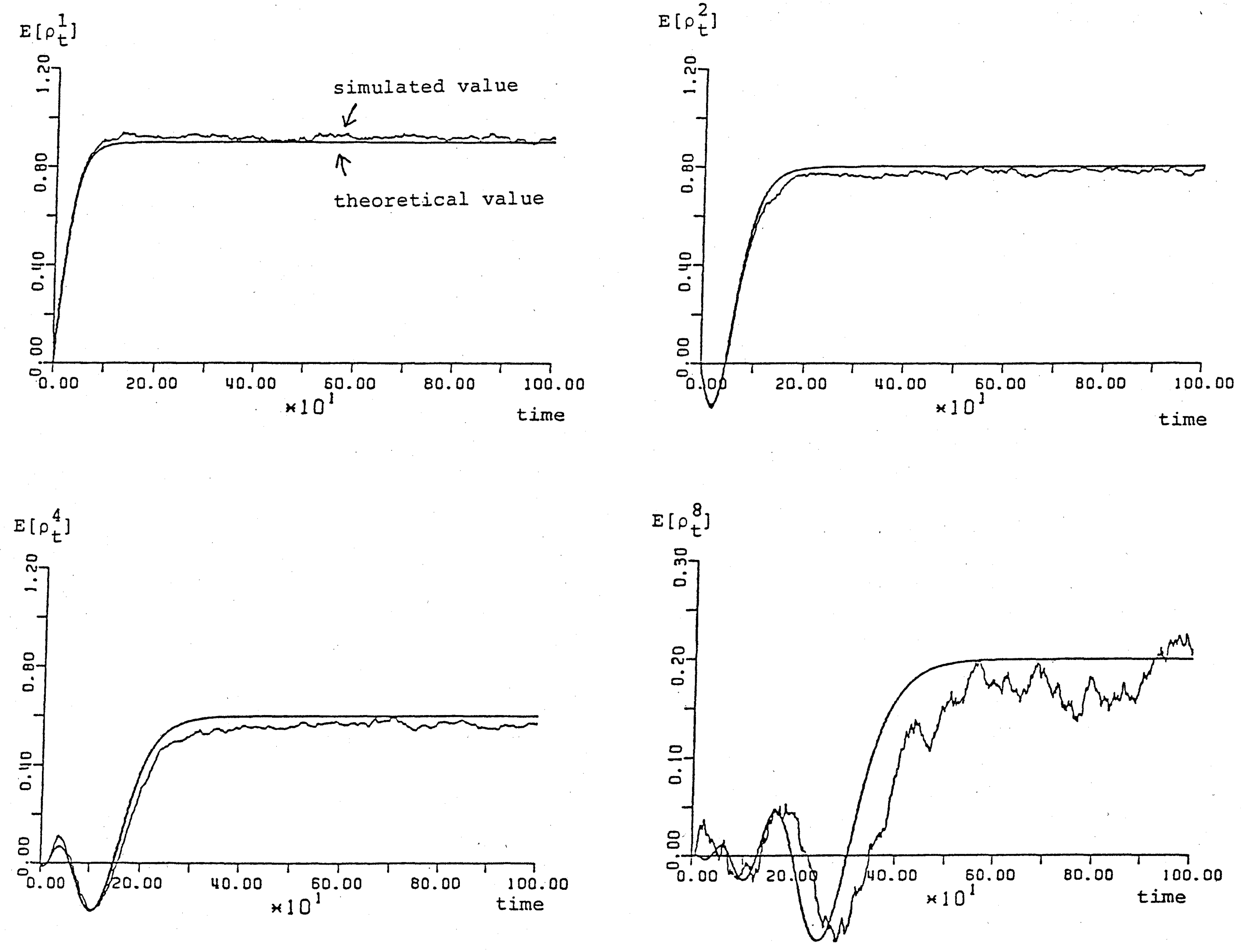

Fig.1 Convergence Models

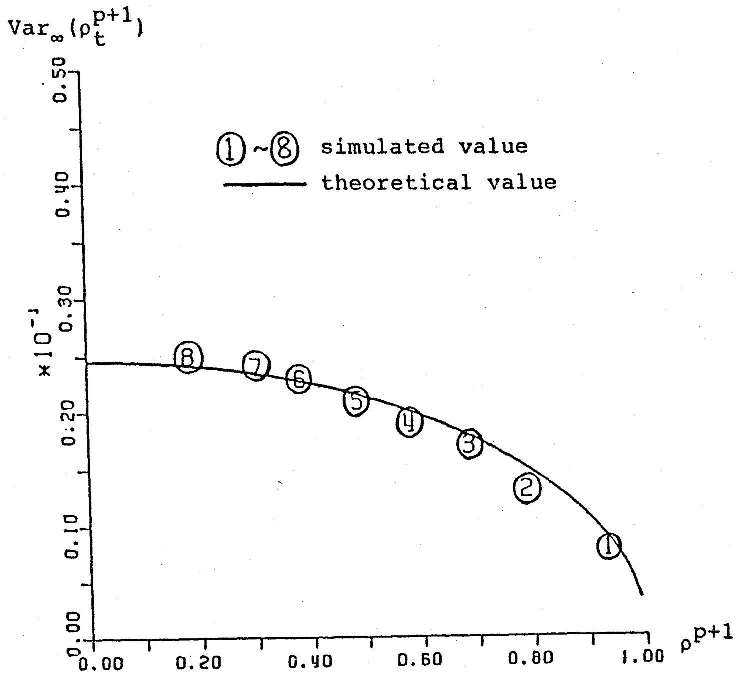

Fig.2 Variance of the estimation error of PARCOR coefficients

Algorithm	μ	$Var_{\infty}(\rho_t^{p+1})$	$\overline{Var}_{\infty}(\rho_t^{p+1})$
A	$\alpha\sigma^2$	$\frac{1}{2}(1-\rho^2)^2\sigma^2\alpha$	$\frac{\mu}{2}(1-\rho^2)^2$
B	$\alpha\sigma^2$	$\frac{1}{2}(1-\rho^2)\sigma^2\alpha$	$\frac{\mu}{2}(1-\rho^2)$
C	$\sqrt{\frac{2}{\pi}}\alpha\sigma$	$\sqrt{\frac{\pi}{8}}(1-\rho^2)\alpha\sigma$	$\frac{\pi}{4}\mu(1-\rho^2)$
D	α	$\frac{\pi^2}{8}(1-\gamma^2)(1-\rho^2)$	$\frac{\pi^2}{8}\mu(1-\gamma^2)(1-\rho^2)$
E	$\frac{2\alpha}{\pi\sqrt{1-\rho^2}}$	$\frac{\pi}{4}\sqrt{1-\rho^2}\,\alpha$	$\frac{\pi^2}{8}\mu(1-\rho^2)$
F	$\sqrt{\frac{2}{\pi}}\frac{\alpha\sigma}{\sqrt{1-\rho^2}}$	$\sqrt{\frac{\pi}{8}}\sqrt{1-\rho^2}\,\alpha\sigma$	$\frac{\pi}{4}\mu(1-\rho^2)$

Table 1 Convergence rates and error variances