(case 1) and Fig. 2 (case 2) for various values of L_d. As can be seen in Fig. 1, interdiffusion initially causes α_T to decrease, followed by an increase, and then a decrease. The initial reduction of α_1 can be explained by the raising of E_1 due to a narrower well width at the bottom of the nonsquare QW for small L_ds, as explained in Reference 7, while E_2 remains at about the same level. This would cause N to be reduced from $4{\cdot}3 \times 10^{17}\,\text{cm}^{-3}$ ($L_d = 0$ Å) to $2{\cdot}1 \times 10^{17}\,\text{cm}^{-3}$ ($L_d = 20$ Å), i.e. a reduction of $\sim 50\%$, and then increase to a maximum of $\sim 10^{18}\,\text{cm}^{-3}$ ($L_d = 40$ Å). However, the α_T peak drops more rapidly with increasing L_d because the negative $\alpha_3(\text{peak}) \propto \mu_{12}^4$ whereas the positive $\alpha_1(\text{peak}) \propto \mu_{12}^2$, where μ_{12} is an increasing function of L_d. A bleaching effect can be seen at large L_ds which is similar to that observed at $L_d = 0$ Å using higher incident optical intensities ($I_{op} > 1.5\,\text{MW/cm}^2$) [4, 7]. The variation of the α_T peak positions, as shown in Fig. 1($a \to f$), is about 153 meV (43 μm) which provides a wide tuning range for the operation of detectors.

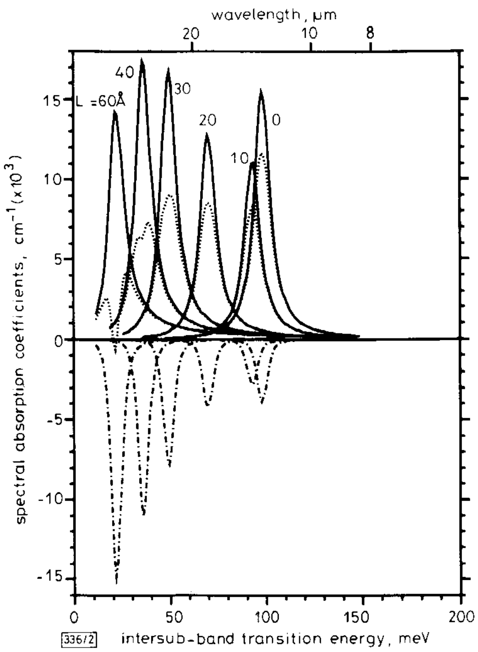

Fig. 2 *Room temperature intersub-band spectral absorption coefficients α_1, α_3, and $\alpha_T = \alpha_1 + \alpha_3$ in 60 Å wide $Al_{0{\cdot}15}Ga_{0{\cdot}85}As/GaAs$ EQW for different diffusion lengths L_d, case 2*

—— α_1
—·— α_3
······ $\alpha_T = \alpha_1 + \alpha_3$

Fig. 2 shows the absorption coefficient for the structure denoted as case 2, which has a lower barrier height and a lower E_f. The average of the absorption peak magnitudes are reduced and the variation of the α_T peak positions is much smaller, 76 meV (13 μm), although the values of α_T are more or less the same. This can provide a more uniform device operation over the corresponding wavelength range. The reduced absorption is due to the lower E_f level causing a smaller N ($5{\cdot}3$–$6{\cdot}3 \times 10^{17}\,\text{cm}^{-3}$), and the small variation of the α_T peaks is due to the small variation of the E_1 levels over the range of L_d used, because it consists of a shallower as-grown QW and therefore the movement of the quantised levels is less sensitive. The bleaching effect is seen to be stronger here.

In summary, we have reported the first theoretical prediction of intersub-band absorption coefficient of interdiffusion induced QWs. The well shape variation can provide a large tuning wavelength range in the far IR region with an almost constant absorption. This may be used to produce a wide bandwidth detector if a series of QWs with different L_ds is used. The advantage of this nonsquare QW structure may also provide lower leakage currents due to reduced tunnelling between the wells, because the barrier thickness at the ground state energy is always thicker than that at the excited state energy.

Acknowledgments: This work is supported by the European Research Office of the US Army and Defence Advanced Research Projects Agency. A. Łaszcz wishes to acknowledge the support of the European Community TEMPUS-JEP.

9th March 1992

E. H. Li, B. L. Weiss and A. Łaszcz (*Department of Electronic and Electrical Engineering, University of Surrey, Guildford, Surrey GU2 5XH, United Kingdom*)

References

1 ZUZZMAN, A., LEVINE, B. F., KUO, J. M., and DEJONG, J.: 'Extended long-wavelength $\lambda = 11$–$15\,\mu$m GaAs/Al$_x$Ga^{1-x}As quantum-well infrared photodetectors', *J. Appl. Phys.*, 1991, **70**, pp. 5101–5107
2 WEST, L. C., and EGLASH, S. J.: 'First observation of an extremely large dipole infrared transition within the conduction band of a GaAs quantum well', *Appl. Phys. Lett.*, 1985, **46**, pp. 1156–1158
3 ROSENCHER, E., and BOIS, PH.: 'Model system for optical nonlinearities: Asymmetric quantum wells', *Phys. Rev. B*, 1991, **44**, pp. 11315–11327
4 KUHN, K. J., IYENGAR, G. U., and YEE, S.: 'Free carrier induced changes in the absorption and refractive index for intersub-band optical transitions in Al$_x$Ga$_{1-x}$As/GaAs/Al$_x$Ga$_{1-x}$As quantum wells', *J. Appl. Phys.*, 1991, **70**, pp. 5010–5017
5 KARUNASIRI, R. P. G., and WANG, K. L.: 'Infrared absorption in parabolic multiquantum well structures', *Superlattices and Microstruct.*, 1988, **4**, pp. 661–664
6 SCHLESINGER, T. E., and KUECH, T.: 'Determination of the interdiffusion of Al and Ga in undoped (Al, Ga)As – GaAs quantum well', *Appl. Phys. Lett.*, 1986, **49**, pp. 519–521
7 AHN, D., and CHUANG, S. L.: 'Calculation of linear and nonlinear intersub-band optical absorptions in a quantum well model with an applied electric field', *IEEE J. Quantum Electron.*, 1987, **QE-23**, pp. 2196–2204
8 COON, D. D., and KARUUASIRI, R. P. G.: 'New mode of IR detection using quantum wells', *App. Phys. Lett.*, 1984, **45**, pp. 649–651

COMPOSITE ALGORITHMS FOR ADAPTIVE IIR FILTERING

S. Lima Netto and P. S. R. Diniz

Indexing terms: Algorithms, Adaptive filtering

Two algorithms based on the combination of the output and equation error schemes are proposed. Their relations are discussed and an example is included to show that the algorithms may be useful for solving some adaptive IIR filtering problems that cannot be successfully solved with existing algorithms.

Introduction: In the parameter estimation of unknown infinite impulse response (IIR) systems the most commonly employed estimation errors are the output error and the equation error [1]. The mean-square equation error (MSEE) surface is well behaved with a single minimum, however equation-error based algorithms converge to biased parameters in the presence of additional noise. On the other hand, the mean-square output error (MSOE) surface may have several local minima depending on the order of the adaptive IIR filter as compared to the unknown system, and on the input signal properties. Output-error based algorithms will converge to a local minimum that may not be acceptable.

Several alternative algorithms are available that combine output and equation error formulations in an attempt to retain the good features of each algorithm [2, 3]. The composite regressor algorithm (CRA) [2] combines the regressor of the equation error with the output error based pseudolinear regressor algorithms [4, 5]. The latter uses a simplified regressor that may stop the algorithm from converging to a minimum of the MSOE surface. Therefore, the CRA converges to biased parameter estimates like the algorithms that originated it.

In this Letter, two algorithms based on the combination of the output and equation error schemes are proposed. Their relations are discussed and an example is included showing that the algorithms may be useful for solving adaptive IIR filtering problems that have not been successfully solved with existing algorithms.

Algorithm I: In a parameter identification problem, the measured output signal of the unknown system is described by

$$y(n) = \left[\frac{B(q^{-1})}{A(q^{-1})}\right]x(n) + v(n) \tag{1}$$

where $B(q^{-1}) = \sum_{j=0}^{n_b} b_j q^{-j}$ and $A(q^{-1}) = 1 + \sum_{i=1}^{n_a} a_i q^{-i}$, are coprime polynomials of the unit delay operator q^{-1}, and $x(n)$ and $v(n)$ are the input signal and the additive noise, respectively. The IIR adaptive filter designed to identify the unknown system is described by

$$\hat{y}(n) = \left[\frac{\hat{B}(q^{-1}, n)}{\hat{A}(q^{-1}, n)}\right]x(n) \tag{2}$$

where $\hat{B}(q^{-1}, n) = \sum_{j=0}^{\hat{n}_b} \hat{b}_j(n)q^{-j}$ and $\hat{A}(q^{-1}, n) = 1 + \sum_{i=1}^{\hat{n}_a} \hat{a}_i(n)q^{-i}$. In output error based algorithms, the objective is to minmise an estimate of the MSOE defined by

$$\begin{aligned} MSOE &= E[e_0^2(n))] \\ &= E\{[y(n) - \hat{y}(n)]^2\} \\ &= E\left\langle\left\{\left[\frac{B(q^{-1})}{A(q^{-1})} - \frac{\hat{B}(q^{-1}, n)}{\hat{A}(q^{-1}, n)}\right]x(n) + v(n)\right\}^2\right\rangle \end{aligned} \tag{3}$$

For equation-error based algorithms, the objective is to minimise an estimate of the MSEE given by

$$\begin{aligned} MSEE &= E[e_E^2(n)] \\ &= E\{[\hat{A}(q^{-1}, n)e_0(n)]^2\} \\ &= E\{[\hat{A}(q^{-1}, n)y(n) - \hat{B}(q^{-1}, n)x(n)]^2\} \end{aligned} \tag{4}$$

The most widely used estimate for the mean square of an arbitrary error $e(n)$ is the instantaneous error squared $e^2(n)$, where $e(n)$ can be the equation, output, or any alternatively defined error. In this case, the updating of the coefficients can be performed as

$$\hat{\theta}(n+1) = \hat{\theta}(n) + \mu e(n)\hat{\phi}(n) \tag{5}$$

where $\hat{\theta}(n) = [\hat{a}_1(n) \ldots \hat{a}_{\hat{n}_a}(n)\hat{b}_0(n) \ldots \hat{b}_{\hat{n}_b}(n)]^T$ is the coefficient vector of the adaptive IIR filter and $\hat{\phi}(n) = -\nabla_\theta[e(n)]$ is the regressor vector, i.e. the gradient vector of $e(n)$ with respect to the coefficient vector.

The CRA combines the equation and output errors by employing an error signal given by

$$e_{CRA}(n) = \left[\frac{\hat{A}(q^{-1}, n)}{\beta + (1-\beta)\hat{A}(q^{-1}, n)}\right]e_0(n) \tag{6}$$

and a regressor vector defined as

$$\hat{\phi}_{CRA}(n) = \beta\hat{\phi}_E(n) + (1-\beta)\hat{\phi}_{PL}(n) \tag{7}$$

where $\hat{\phi}_E(n) = [y(n-1) \ldots y(n-\hat{n}_a)x(n) \ldots x(n-\hat{n}_b)]^T$ is the equation error algorithm regressor $\hat{\phi}_{PL}(n) = [\hat{y}(n-1) \ldots \hat{y}(n-\hat{n}_a)x(n) \ldots x(n-\hat{n}_b)]^T$ is the regressor of the pseudolinear algorithm, and β is a composition factor. It can be noticed that the regressor used in the CRA does not correspond to the gradient of $e_{CRA}(n)$ with respect to the parameters, i.e. the CRA does not have the basic property of minimising an estimate of $E[e_{CRA}^2(n)]$ following the steepest-descent direction.

An algorithm based on the error $e_{CRA}(n)$, denoted by algorithm I, can be derived by using the regressor defined by

$$\begin{aligned} \hat{\phi}_I(n) &= -\frac{\partial e_{CRA}(n)}{\partial\hat{\theta}} \\ &= \frac{1}{[\beta + (1-\beta)\hat{A}(q^{-1}, n)]^2} \\ &\quad \times \begin{Bmatrix}\beta y(n-i) + [(1-\beta)\hat{A}(q^{-1}, n)]\hat{y}(n-i) \\ \beta x(n-j) + [(1-\beta)\hat{A}(q^{-1}, n)]x(n-j)\end{Bmatrix} \\ &= \frac{1}{[\beta + (1-\beta)\hat{A}(q^{-1}, n)]^2} \\ &\quad \times [\beta\hat{\phi}_E(n) + (1-\beta)\hat{A}(q^{-1}, n)\hat{\phi}_{PL}(n)] \end{aligned} \tag{8}$$

The coefficient updating of algorithm I is then described by

$$\hat{\theta}(n+1) = \hat{\theta}(n) + \mu e_I(n)\hat{\phi}_I(n) \tag{9}$$

where $e_I(n) = e_{CRA}(n)$ and $\hat{\phi}_I(n)$ are defined, respectively, in eqns. 6 and 8.

Algorithm II: Algorithm II is derived by combining the output and equation error in the following form:

$$\begin{aligned} e_{II}(n) &= \beta e_E(n) + (1-\beta)e_0(n) \\ &= \{1 + \beta[\hat{A}(q^{-1}, n) - 1]\}e_0(n) \end{aligned} \tag{10}$$

with the regressor vector given by

$$\begin{aligned} \hat{\phi}_{II}(n) &= -\frac{\partial e_{II}(n)}{\partial\hat{\theta}} \\ &= \begin{Bmatrix}\beta y(n-i) + \left[\frac{(1-\beta)}{\hat{A}(q^{-1}, n)}\right]\hat{y}(n-i) \\ \beta x(n-j) + \left[\frac{(1-\beta)}{\hat{A}(q^{-1}, n)}\right]x(n-j)\end{Bmatrix} \\ &= [\beta\hat{\phi}_E(n) + (1-\beta)\hat{\phi}_0(n)] \end{aligned} \tag{11}$$

The coefficient updating is performed as follows:

$$\hat{\theta}(n+1) = \hat{\theta}(n) + \mu e_{II}(n)\hat{\phi}_{II}(n) \tag{12}$$

From eqns. 6, 8, 10, and 11 the following characteristics of the algorithms I and II can be easily shown:

$$e_I(n)\Big|_{\beta=0} = e_{II}(n)\Big|_{\beta=0} = e_0(n) \tag{13}$$

$$e_I(n)\Big|_{\beta=1} = e_{II}(n)\Big|_{\beta=1} = e_E(n) \tag{14}$$

$$\begin{aligned} \hat{\phi}_I(n)\Big|_{\beta=0} &= \frac{\partial e_I(n)}{\partial\hat{\theta}}\Big|_{\beta=0} = \hat{\phi}_{II}(n)\Big|_{\beta=0} \\ &= \frac{\partial e_{II}(n)}{\partial\hat{\theta}}\Big|_{\beta=0} = \hat{\phi}_0(n) \end{aligned} \tag{15}$$

$$\begin{aligned} \hat{\phi}_I(n)\Big|_{\beta=1} &= \frac{\partial e_I(n)}{\partial\hat{\theta}}\Big|_{\beta=1} = \hat{\phi}_{II}(n)\Big|_{\beta=1} \\ &= \frac{\partial e_{II}(n)}{\partial\hat{\theta}}\Big|_{\beta=1} = \hat{\phi}_E(n) \end{aligned} \tag{16}$$

In fact, the errors in algorithms I and II are related by

$$e_{II}(n) = e_I(n) + \left\{\frac{\beta(1-\beta)[\hat{A}(q^{-1}, n) - 1]^2}{\beta + (1-\beta)\hat{A}(q^{-1}, n)}\right\}e_0(n) \tag{17}$$

Despite this relation, algorithms I and II might have distinct overall properties, for example algorithm I has higher computational complexity than algorithm II. For this reason, only algorithm II is considered in the following Section.

Simulation results: Algorithm II was tested for a number of system identification problems. Here we concentrate on the identification of a second-order plant with transfer function [7]

$$H(z^{-1}) = \frac{0.05 - 0.4z^{-1}}{1 - 1.1314z^{-1} + 0.25z^{-2}} \quad (18)$$

where the adaptive filter is a first-order filter described by

$$H(z^{-1}, n) = \frac{\hat{b}_0(n)}{1 - \hat{a}_1(n)z^{-1}} \quad (19)$$

The input signal is a white noise with zero mean and unit variance. Given the initial point at $[\hat{b}_0(0); \hat{a}_1(0)] = [-0.1; 0.2]$, it was found that algorithm II with $\beta = 0.04$ converges to the point $[-0.306; 0.907]$ that is closer to the global minimum $[-0.311; 0.906]$ than the equation error algorithm solution $[0.050; 0.931]$. The output error algorithm converges to a local minimum $[0.114; -0.519]$. In fact, the error surface for $\beta = 0.04$ is unimodal as depicted in Fig. 1. For $\beta = 0.03$ a local minimum still exists. A strategy that can be used to reach the global minimum is to start with $\beta = 1$, that corresponds to the equation error criteria, and decrease β as the algorithm progresses. This approach was implemented in the example above reducing β by 0·1 every 1000 iterations, and as a result the global minimum was reached.

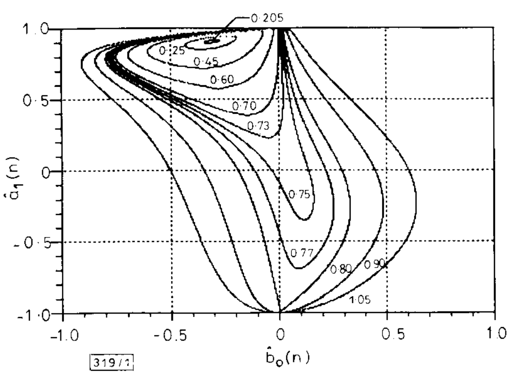

Fig. 1 *Error surface contours for $\beta = 0.04$*

It should be mentioned that the proposed algorithms do not guarantee convergence to global minimum of the MSOE surface, however the solution may be better than the one obtained with the equation error algorithm for insufficient order identification, or in the presence of additional noise.

Conclusion: This work presented alternative algorithms for adaptive IIR filtering based on appropriate combination of the output and equation error approaches. The proposed algorithms provide an extra parameter that may be manipulated to obtain a tradeoff between parameter bias and convergence to a local minimum. Further studies are needed to device a strategy to implement a variable β.

9th March 1992

S. Lima Netto and P. S. R. Diniz* (*Programa de Engenharia Elétrica, COPPE/EE/Federal University of Rio de Janeiro, Caixa Postal 68504, Rio de Janeiro, RJ 21945, Brazil*)

* Mailing address: Department of Electrical and Computer Engineering, University of Victoria, PO Box 3055, Victoria BC V8N 3P6, Canada

References

1 SHYNK, J. J.: 'Adaptive IIR filtering', *IEEE Acoust., Speech, Signal Processing Magazine*, 1989, **6**, (2), pp. 4–21
2 KENNEY, J. B., and ROHRS, C. E.: 'The composite regressor algorithm'. Proc. Int. Conf. Acoust., Speech, Signal Processing, New York, April 1989, pp. 1561–1563
3 GERALD, J. A. B., ESTEVES, N. L., and SILVA, M. M.: 'An adaptive IIR echo canceller combining output-error criteria'. Proc. Int. Symp. Circuits Syst., New Orleans, May 1990, pp. 779–783
4 FEINTUCH, P. L.: 'An adaptive recursive LMS filter', *Proc. IEEE*, 1976, **64**, (11), pp. 1622–1624
5 LANDAU, I. D.: 'Unbiased recursive identification using model reference adaptive techniques', *IEEE Trans.*, 1976, **AC-21**, (2), pp. 194–202
6 GEE, S., and RUPP, M.: 'A comparison of adaptive IIR echo canceller hybrids'. Proc. Int. Conf. Acoust., Speech, Signal Processing, Toronto, Canada, May 1991, pp. 1541–1544
7 JOHNSON, C. R., JUN., and LARIMORE, M. G.: 'Comments on and additions to "An adaptive recursive LMS filter"', *Proc. IEEE*, 1977, **65**, (9), pp. 1399–1402

MODIFIED EXPONENTIAL BIDIRECTIONAL ASSOCIATIVE MEMORIES

W.-J. Wang and D.-L. Lee

Indexing terms: Associative memories, Neural networks

Based on the Jeng exponential bidirectional associative memory (EBAM), a modified updating rule of EBAM is presented. In the recalling processes of the modified EBAM (MEBAM), the continuity assumption of the EBAM is relaxed and heterocorrelation processes run parallel with the autocorrelation processes. An energy function, which is defined and does not increase on the change of neuron states, ensures the stability of the system. Finally computer simulations demonstrates that the MEBAM has much better storage capacity than that of the Jeng EBAM.

Introduction: The directional associative memory [1–3] (BAM) is a minimal two-layer nonlinear feedback network in which associative paired data are recalled by directionally updating the neuron state through the connection matrix M and its tranpose M^T. The primary constraint is that all the stored pairs are local minima of the energy surface. Simpson [4] proposed the intraconnected BAM (IBAM) by adding the intralayer connections to the Kosko BAM, with which the complement encoding problem can be removed. However, the continuity assumption of the BAM had not been relaxed until the modified IBAM was proposed [7]. Furthermore, Jeng *et al.* [6] proposed the exponential BAM (EBAM) which uses exponential nonlinearity to improve the storage capacity of the BAM. However, only heterocorrelation updating processes exist in the recall process of the EBAM which still requires the continuity assumption. We propose a modified updating rule for the EBAM by adding an autocorrelation exponential term to the original phase of the EBAM such that the continuity assumption is relaxed and the storage capacity is improved. An energy function which is defined and does not increase during the change of the neuron states ensures the stability of the system.

Analysis of EBAM: Assume there are m stored pairs $\{(X^i, Y^i)\}_{i=1}^{m}$, where $X^i \in \{-1, 1\}^n$ and $Y^i \in \{-1, 1\}^p$. The bidirectional updating rules [6] of the EBAM in one cycle are

$$Y = f(X) = \operatorname{sgn}\left\{\sum_{i=1}^{m} Y^i \alpha^{\langle X^i, X\rangle}\right\} \quad (1a)$$

$$X' = g(Y) = \operatorname{sgn}\left\{\sum_{i=1}^{m} X^i \alpha^{\langle Y^i, Y\rangle}\right\} \quad (1b)$$

where $\alpha > 1$ and $\langle\ ,\ \rangle$ denotes the inner product; sign $(x) = 1$ for $x \geq 0$ and sgn $(x) = -1$ otherwise. As mentioned by Jeng *et al.* [6], the EBAM has improved the storage capacity and the error correcting capability of the BAM. However, the 'continuity assumption' of the Kosko BAM [3]

$$\frac{1}{n} H(X^i, X^j) \simeq \frac{1}{p} H(Y^i, Y^j) \quad (2)$$