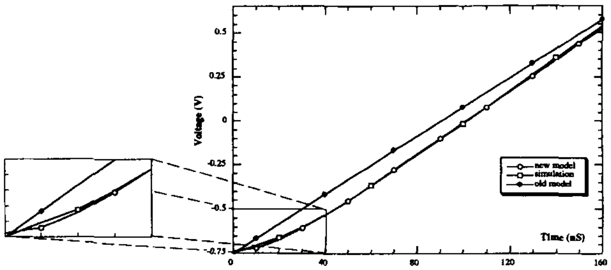


Fig. 4. Slew rate behavior: Simulation versus model.

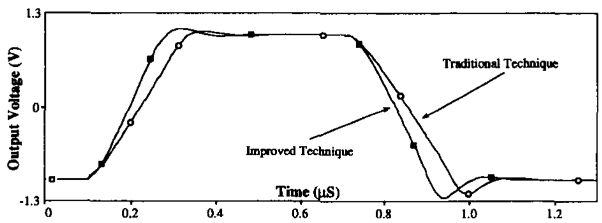


Fig. 5. The step response for the traditional and improved designs.

that the deviation is largest for small input steps as is predicted by (14) and (15). The traditional simplified model is only able to approximate the slewing behavior for large phase margin and a small slew rate/bandwidth ratio. Our model on the other hand is general and can predict the slewing behavior for all conditions.

We have designed two circuits, one using the simplified expression for slew rate and the other using our model for slew rate. Both circuits were designed for ($SR = 10$ V/μS, $C_L = 10$ pF, $C_c = 10$ pF, $UGF = 6$ MHz, $\phi_m = 45°$). Other design parameters and measured and predicted slew rates are shown in the bottom half of Table I. The numbers in top row of the bottom half correspond to the circuit designed using the simplified model and the numbers in the bottom row correspond to the circuit designed using our model. The slew rate numbers in the last two columns are values predicted by our model. It is clear that more predictable slew rates are possible by using our model. The output waveform for a positive and a negative step for the two circuits is shown in Fig. 5.

IV. Conclusion

We have analyzed the slewing behavior of a two stage Miller compensated CMOS amplifier. In this paper we have presented analytical expressions for the amplifier response during slewing. Using this expression for the output voltage we have developed an improved formula to calculate the effective slew rate. We also presented a closed form equation to determine the necessary tail current. The effect of the load capacitance, device sizes and bias currents are included in these new models. It was shown that small-signal models cannot be used to predict slewing behavior. Results from SPICE simulations were used to validate our analysis and design methodology. Our model is general and is applicable for all design parameters and simplifies to the traditional model ($SR = I_1/C_c$) for small values of γ.

References

[1] F. Wang and R. Harjani, "Dynamic amplifiers: Settling, slewing and power issues," *IEEE Int. Symp. Circuits Syst.*, 1995.

[2] C. T. Chuang, "Analysis of the settling behavior of an operational amplifier," *IEEE J. Solid-State Circuits*, vol. SC-17, pp. 74–80, Feb. 1982.

[3] J. C. Lin and J. H. Nevin, "A modified time-domain model for nonlinear analysis of an operational amplifier," *IEEE J. Solid-State Circuits*, vol. SC-21, pp. 478–483, June 1986.

[4] B. Y. Kamath, R. G. Meyer, and P. R. Gray, "Relationship between frequency response and settling time of operational amplifiers," *IEEE J. Solid-State Circuits*, vol. SC-9, pp. 347–352, Dec. 1974.

A New Composite Gradient Algorithm to Achieve Global Convergence

Gyula Simon and Gábor Péceli

***Abstract*—Insufficient-order system identification can result in a multimodal mean square error surface on which a gradient-type algorithm may converge to a local minimum. In this letter a new composite gradient algorithm (CGA) is presented which is due to achieve global convergence when the output error surface contains local minima. The proposed algorithm combines the useful properties of the output error (OE) and equation error (EE) adaptive filtering methods using a new dynamic error surface. The CGA provides a single convergence point for the gradient-search algorithm independently of the initial conditions. The "global convergence" conjecture is illustrated by simulation examples showing good global convergence properties even in such undermodeled cases when the Steiglitz-McBride algorithm fails.**

I. Introduction

Due to their simplicity gradient search methods play an important role in recursive identification and adaptive filtering algorithms. One of the best known gradient-type algorithm is the output error (OE) method which finds the minimum of the mean square error (MSE) surface related to the output error. The OE algorithm performs well when the error surface is unimodal. In practical cases, however, often there is no *a priori* knowledge of the degree of the unknown system or it is too large. In these cases the model is "insufficient" and the error surface may be multimodal [1]–[3].

There are known conditions for the unimodality of the OE surface [1], [2], but since the system is unknown, these requirements often cannot be met. Thus different initial points can lead the algorithm to different (local) minima. Another possible way to overcome the problem of local minima is the use of other error surfaces. The equation error (EE) surface is a unimodal (quadratic) function of the coefficients but the minimum point can be far from the global minimum of the OE surface [1]. Promising solutions are the algorithms with error surfaces composed from "OE-like" and "EE-like" elements [4], [5]. In the next chapter some known algorithms are surveyed and then a new dynamic composite algorithm is proposed with illustrative properties of the error surface.

II. Gradient Algorithms with Composite Error Surfaces

Fig. 1 shows a general adaptive filter model with the unknown system $H(z)$, the input, desired and the output signals are $x(n)$,

Manuscript received August 3, 1994; revised June 3, 1995. This paper was recommended by Associate Editor S. Kiaei.

The authors are with the Department of Measurement Engineering, Technical University of Budapest, Budapest, Hungary.

IEEE Log Number 9414319.

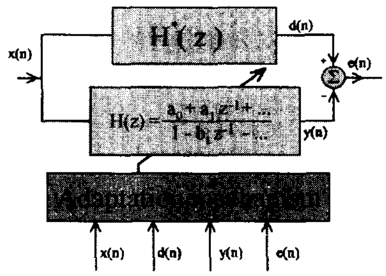


Fig. 1. Adaptive system model.

$d(n)$ and $y(n)$, respectively. We will assume here the input signal to be white (in the experiments normally distributed random noise was used). The adaptation mechanism gets $x(n), d(n)$ and $e(n)$ and tries to adjust the filter coefficients α's and β's so that the mean square error (i.e., $E\{e^2\}$) is minimum.

When local minima exist a simple OE gradient search method may converge to a shallow local minimum point instead of the global one. Fig. 2(a) shows an OE surface with two minimum points [3].

It is clear that a simple gradient search method cannot guarantee global convergence on a surface like this. A possible solution can be the use of a composite error surface (CES) on which the gradient search leads the parameters to or close to the global minimum point of the original OE surface.

The simplest CES is the EE surface. In this case the adaptation mechanism contains an EE algorithm, and the OE model parameters are the copies of the parameters of the EE model. This method provides one convergence point but since the minimum point of EE method usually differs from that of the OE method, this algorithm generally has bias as shown in Fig. 2(b).

A good example of CES is the Steiglitz-McBride method. The SMM defines the gradient of the CES the following way:

$$\mathrm{Grad} = E\left\{e_{\mathrm{OE}} \cdot \frac{1}{1 - A(z^{-1})}\begin{bmatrix} x(n) \\ d(n) \end{bmatrix}\right\}. \tag{1}$$

The difference between the SMM and the OE method is that in the expression of gradient the output signal $y(n)$ is replaced by the desired signal $d(n)$, which renders the properties of the algorithm between the EE and OE methods. Fig. 2(c) shows an example of the CES of the SMM. The example clearly shows that the CES of the SMM has one minimum point and it is close to the global minimum point of the OE surface. This "global convergence" phenomenon was observed and partly explained in [4], and necessary conditions are also presented. In many practical cases these requirements are met, but there can be one or more minima which can be far from either the global or local minima of the OE surface (see Fig. 5).

The Composite Regressor Algorithm (CRA) [5] can be considered as a dynamic CES method, although in [5] the multimodal case is not considered and the structure is not completely equivalent to the model shown in Fig. 1. The regressor is defined as a convex sum of an OE and EE regressor

$$\begin{bmatrix} x(n) \\ y(n) \end{bmatrix}_{\mathrm{CRA}} = (1 - \gamma(n))\begin{bmatrix} x(n) \\ y(n) \end{bmatrix} + \gamma(n)\begin{bmatrix} x(n) \\ d(n) \end{bmatrix}. \tag{2}$$

The composition of the regressor—and thus the error surface—is determined by the regression composition parameter $\gamma(n)$, which can vary in time. If $\gamma = 1$, the CRA is equivalent to the EE method, while $\gamma = 0$ means a pure OE algorithm.

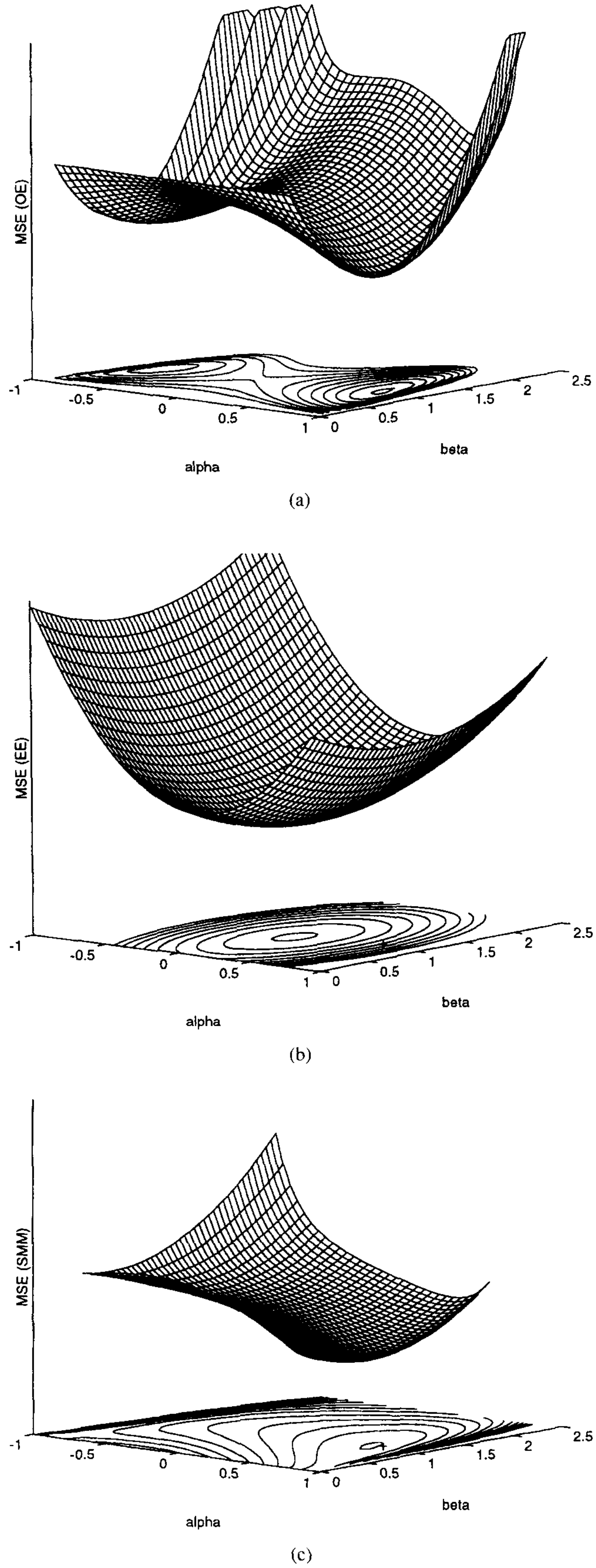


Fig. 2. (a) OE surface, (b) EE surface, (c) SMM surface. $H(z^{-1}) = 1 + 0.25z^{-1} + 2z^{-2}$, $B(z^{-1}) = 3$, $A(z^{-1}) = 1 - \alpha$.

The proposed Composite Gradient Algorithm is the based upon a similar idea.

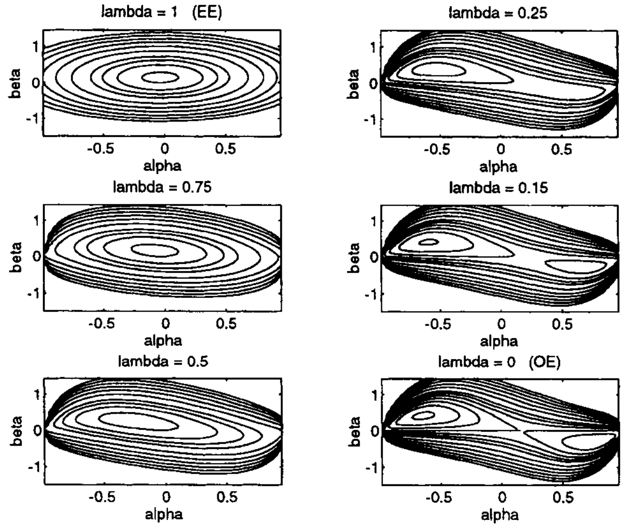


Fig. 3. Composite error surface of CGA versus composition parameter.

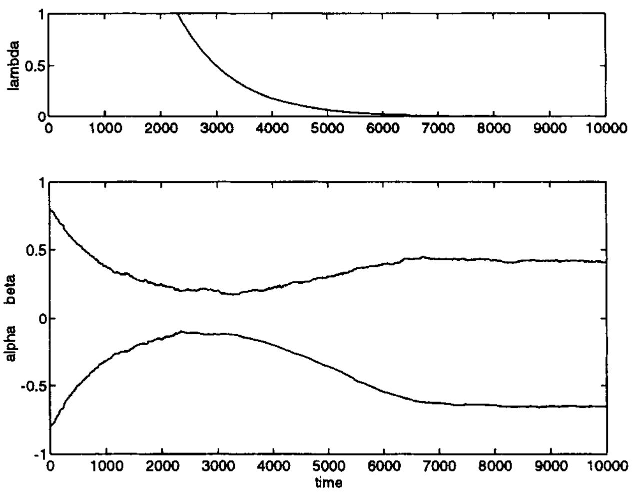


Fig. 4. Composition parameter and filter parameters versus time.

III. The Composite Gradient Algorithm (CGA)

Consider the adaptive system shown in Fig. 1. The output of the filter and the (output) error are

$$y(n) = y_{\mathrm{OE}}(n) = \frac{B(z^{-1})}{1 - A(z^{-1})} x(n),$$
$$e(n) = e_{\mathrm{OE}}(n) = d(n) - y(n). \tag{3}$$

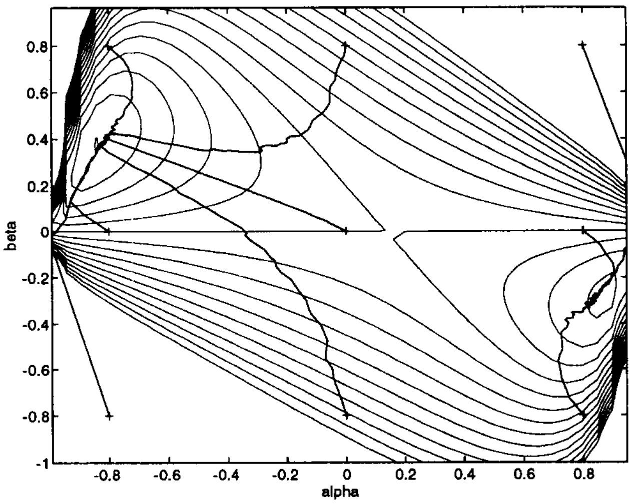
The adaptation mechanism contains an EE structure with output y_e and error e_e as follows:

$$y_{\mathrm{EE}}(n) = B(z^{-1})x(n) + A(z^{-1})d(n),$$
$$e_{\mathrm{EE}}(n) = d(n) - y_{\mathrm{EE}}(n). \tag{4}$$

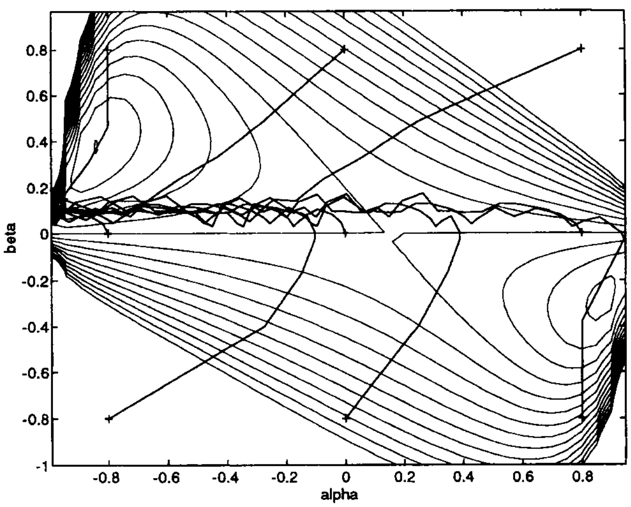
The parameters α's and β's of the EE structure are the same as in (3). Let us define a MSE surface as the convex sum of the OE and EE MSE surfaces defined by (3) and (4). Thus the gradient of this CES can be computed as the convex sum of the OE and EE gradients

$$\begin{aligned}\mathrm{Grad}_{\mathrm{CGA}} &= (1-\lambda)\mathrm{Grad}_{\mathrm{OE}} + \lambda \mathrm{Grad}_{\mathrm{EE}} \\ &= -2(1-\lambda)e_{\mathrm{OE}} \frac{1}{1 - A(z^{-1})} \begin{bmatrix} x(n) \\ y_{\mathrm{OE}}(n) \end{bmatrix} \\ &\quad - 2\lambda e_{\mathrm{EE}} \begin{bmatrix} x(n) \\ d(n) \end{bmatrix}\end{aligned} \tag{5}$$

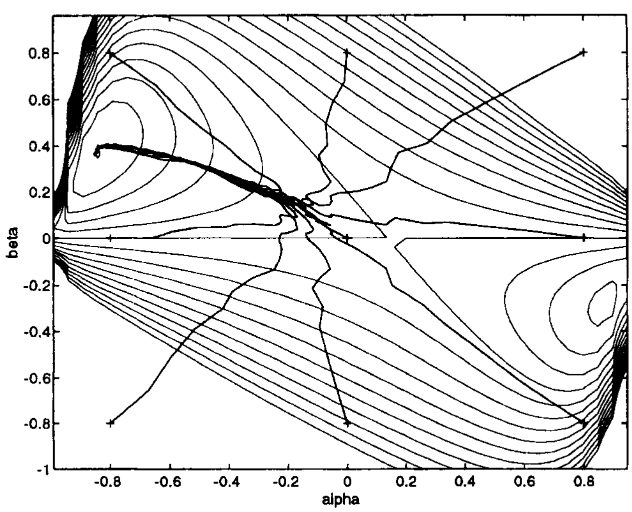
where λ is the gradient composition parameter.



(a)



(b)



(c)

Fig. 5. (a) Parameter trajectories of SMM from different initial points. (b) Parameter trajectories of CRA from different initial points. (c) Parameter trajectories of CGA from different initial points.

The parameter update mechanism is a simple gradient search method

$$\begin{bmatrix} \underline{\beta}(n+1) \\ \underline{\alpha}(n+1) \end{bmatrix} = \begin{bmatrix} \underline{\beta}(n) \\ \underline{\alpha}(n) \end{bmatrix} - \delta \cdot \mathrm{Grad}_{\mathrm{CGA}} \tag{6}$$

where δ is the stepsize.

Fig. 3 shows an example how the distorted error surface of the CGA changes from EE to OE MSE surface vs. the composition parameter. It is clearly visible that the original (EE) global minimum moves to the global minimum of the OE surface as λ decreases and the local minimum appears only later. The CGA algorithm starts with $\lambda = 1$ and waits until the parameters converge to the (EE) minimum. After convergence to the EE minimum the composition parameter is decreased and the parameters trapped by the pitfall move along with the new error surface in the wandering "global" minimum until the composition parameter decreases to zero and the error surface becomes the OE surface. Once the parameters are trapped by a minimum (hopefully by the "global" one), the other minima appearing later cannot affect the convergence and the parameters are led to the global minimum of the OE surface.

IV. Convergence Properties of the CGA

While the composition parameter λ is close to 1 the behavior of the algorithm is identical to that of the EE, so the speed of convergence is determined by the eigenvalue spread of the Hessian matrix [1]. If λ tends to zero the error surface becomes similar to the OE surface where the convergence analysis is largely unsolved, but the speed of convergence is probably influenced by the eigenvalues of the Hessian [1].

The composition parameter must be set to 1 until the parameters find the original minimum of the EE surface. Then the decay rate of λ must be "small enough" so that the parameters can follow the wandering minimum. In the next examples the composition parameter update is based upon a slowly decreasing exponential. A possible monitoring method could be based upon the estimation of the eigenvalue spread of the Hessian.

The necessary conditions for global convergence are not known yet, but the algorithm showed promising behavior for low order cases in exhaustive simulation tests. During the test a fixed filter with maximum two zeros or two poles and an adaptive filter with one pole and zero were used and the CGA always converged to the global minimum. In the next chapter some simulation examples illustrate the "global convergence" property of the CGA.

V. Simulation Examples

The algorithm is defined by (3)–(6). A second order "unknown" system and a first order adaptive filter were used to provide multimodal MSE

$$H(z^{-1}) = \frac{b_0 + b_1 z^{-1}}{1 - a_0 z^{-1} - a_1 z^{-2}}, \frac{B(z^{-1})}{1 - A(z^{-1})} = \frac{\beta}{1 - \alpha z^{-1}}. \tag{7}$$

Fig. 4 shows a simulation example with the same system as shown in Fig. 3. The parameters are $b = [0.1248, -0.9981]$ and $a = [0.1128, -.0036]$. The plots indicate the composition parameter and the filter coefficients during convergence. It is clearly visible that the parameters first converge to the EE minimum and then find the global minimum of the OE surface.

The system of Fig. 5 ($b = [0.0925, -0.7398], a = [0.0303, 0.6642]$) does not satisfy the requirements stated in [4], thus the SMM failed to converge to the global minimum when the initial parameters were close to the local minimum. The CRA does not use true gradient values [5] so the algorithm can not find the true global minimum. The CGA performed well independently of the starting point.

VI. Conclusion

In this letter a new candidate algorithm was presented to achieve global convergence in case of multimodal mean square error surface. The algorithm uses a composite error surface composed from OE and EE surfaces and the global convergence is achieved by dynamic update of the composition. The algorithm is known to have one possible minimum without bias. The global convergence and stability properties of the algorithm are under investigation, but the simulation examples are promising.

References

[1] J. J. Shynk, "Adaptive IIR filtering," *IEEE ASSP Mag.*, vol. 6, pp. 4–21, Apr. 1989.

[2] T. Söderström and P. Stoica, "Some properties of the output error method," *Automat.*, vol. 18, pp. 93–99, Jan. 1982.

[3] S. D. Stearns, "Error surfaces of recursive adaptive filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-29, pp. 763–766, June 1981.

[4] H. Fan and M. Doroslovački, "On 'global convergence' of Steiglitz-McBride adaptive algorithm," *IEEE Trans. Circuits Syst. II*, vol. 40, pp. 73–87, Feb. 1993.

[5] J. B. Kenney and C. E. Rohrs, "The composite regressor algorithm for IIR adaptive systems," *IEEE Trans. Signal Process.*, vol. 41, pp. 617–628, Feb. 1993.