

BLIND EQUALIZATION SCHEMES WITH DIFFERENT ERROR EQUATIONS

Pulakesh Roy and A. A. (Louis) Beex

Systems Group – DSP Research Laboratory
The Bradley Department of Electrical and Computer Engineering
Virginia Tech, Blacksburg, VA 24061– 0111

ABSTRACT

A general problem with blind adaptation techniques is that they have poor convergence properties compared to the traditional techniques using training sequences. Having a multi-modal cost surface, blind adaptation techniques may force the equalizer to converge to a false minimum, depending on the initialization. The most commonly used blind adaptation algorithm is the *Constant Modulus Algorithm* (CMA). In our simulations a logarithmic error equation causes CMA to converge to a global minimum when a differential encoding scheme is used. The convergence rates of CMA with different error equations were also investigated for a symmetric channel.

1. INTRODUCTION

The model of the equalizer used for the error analysis is shown in Figure 1. In a blind equalization scheme, instead of using a training sequence, some statistical property of the signal is used for the determination of the instantaneous error, $e(n)$, which is then used for updating the equalizer coefficient vector $\mathbf{f}(n)$.

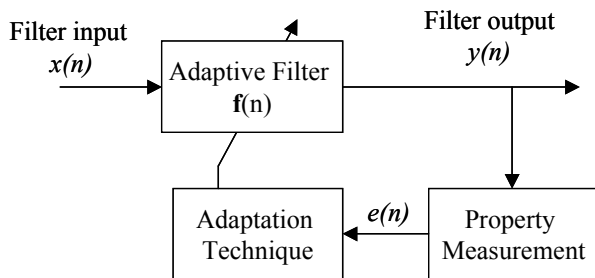


Figure 1: Adaptive Filter Diagram for Blind Equalization.

Generally, the constant modularity of the output signal $y(n)$ is the desired property that is to be restored. The adaptive algorithm that works mainly to restore the constant modulus property of the output signal is known as the *Constant Modulus Algorithm* (CMA), and the same algorithm will be used to analyze the properties of different error equations. Error equations used in the *property measurement* block of Figure 1 measure the deviation of the output signal amplitude from the desired amplitude level and force the adaptive algorithm to change the equalizer coefficients to compensate for the present error. To what extent the adaptive algorithm compensates for a specific deviation of the output signal

amplitude from the desired amplitude level is completely dependent on the error equation. Consequently the error equation used in an adaptive system exercises significant control over the overall system performance.

The convergence properties of any adaptive algorithm depend on the cost function, which is subject to minimization during the adaptation process. The cost function is a function of the equation for the error, defined as the difference between the present and the desired value of any property of the signal that is to be restored. Therefore, the cost function of an adaptive algorithm can be changed either by changing the function itself or by changing the error equation. The most commonly used definition of the cost function is the mean squared value of the error. It has been shown [2] that changing the definition of the cost function provides a lot of advantages, especially in the time varying situation. In this work, instead of changing the definition of the cost function, different types of error equations are used to change the CM cost surface. In the following sections three different types of error equation are defined and analyzed.

2. ERROR SIGNALS

Out of the many possible error equations the following are selected for analysis. The equations were suggested earlier [1].

$$\begin{aligned} \text{Type 1: } e(n) &= \frac{1}{2A} \left\{ |y(n)|^2 - A^2 \right\} \\ \text{Type 2: } e(n) &= |y(n)| - A \\ \text{Type 3: } e(n) &= A \ln \left(\frac{|y(n)|}{A} \right) \end{aligned} \quad (1)$$

In (1), A is the desired amplitude level. The error equations are non-linear and this non-linearity causes the blind equalization scheme to be non-linear. All the error equations become zero when the amplitude of $y(n)$ has the desired value. Moreover the error equations have the same slope at the zero error point. They also provide sign information to indicate whether the absolute value of $y(n)$ is less than or greater than the desired value A . This sign information determines the direction of adaptation. Figure 2 shows how the error varies with the absolute value of the filter output $y(n)$ for different error types. A

desired output-amplitude of $A=1$ was used in plotting the error curves appearing below.

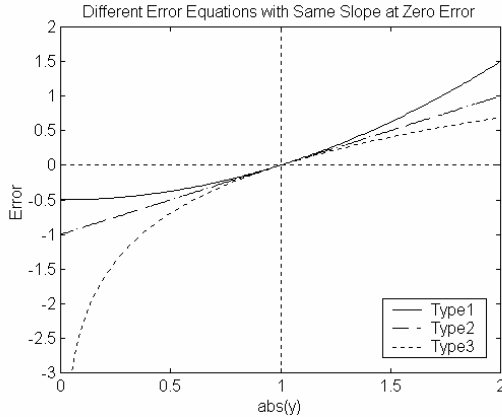


Figure 2: Types of Error Functions Analyzed.

Type 1 is the most commonly used error equation with CMA. The adaptive algorithm with a Type 1 error equation compensates more when the amplitude level of the output signal is higher than the desired amplitude level. This kind of compensation usually results in a better convergence rate at the beginning of the adaptation process since, generally, in this situation the error signal has a large positive value. However, near the convergence point, or the zero error point, the Type 1 error is not symmetric. Due to this non-symmetric behavior, near the convergence point, for the same positive and negative deviations of the amplitude level from its desired value, the Type 1 error does not provide the same compensation. Therefore, the Type 1 error is expected to provide a higher steady state error.

Since an affine relationship exists between the Type 2 error and the equalizer output amplitude, Type 2 is also referred to as the *affine error*. The error equation of Type 2 is symmetric about zero near the minimum of the cost surface. Therefore with this error equation type, the steady state performance of CMA is expected to be better than for Type 1. The convergence rate with a Type 2 error equation may not be as good as for Type 1 at the beginning of the adaptation process if the amplitude of the equalizer output has a much higher value than the desired amplitude level. For the same reason, a Type 3 error is expected to yield worse performance at the beginning of the adaptation process.

The error equation determines the equations for updating the adaptive filter coefficients (or tap weights). Therefore, the weight update equations for different types of error are derived in the next section. Again, for any of the above types of error, the system becomes non-linear and analytically not very tractable. Therefore simulation results are used to compare the convergence behavior with different types of error equation.

3. WEIGHT UPDATE EQUATIONS

The most commonly used cost function $J[e(n)]$ is the expected value of the square of the error signal. Updating the coefficients of the filter is aimed at optimization of the cost function, i.e. minimizing the mean squared error. The same cost function will be used to derive the weight update equations.

$$J[e(n)] = E\{e^2(n)\} \quad (2)$$

$E\{\bullet\}$ in (2) is the statistical expectation operator. Thus $J[e(n)]$ is the mean square of the error sensed in measuring adherence to the constant modulus property. The constant modulus algorithm (CMA) employs an approximate gradient descent method to minimize the cost function given in (2), and the update equation for the filter coefficients becomes,

$$\mathbf{f}(n+1) = \mathbf{f}(n) - \mu \nabla_{\mathbf{r}} J \quad (3)$$

In (3), $\mathbf{f}(n)$ is the tap weight vector at instant n , and μ is the adaptation step size, a small positive scalar. The true gradient in (3) is approximated by its instantaneous value $\hat{\nabla}_{\mathbf{r}} J$.

$$\hat{\nabla}_{\mathbf{r}} J = \frac{\partial}{\partial \mathbf{f}(n)} \{e^2(n)\} = 2e(n) \frac{\partial e(n)}{\partial \mathbf{f}(n)} \quad (4)$$

The specific weight update equations corresponding to (3) and (4) are now derived for the different types of error given in (1).

3.1 Type 1 Error

$$e(n) = \left\{ |y(n)|^2 - A^2 \right\} / 2A = \left\{ y^*(n)y(n) - A^2 \right\} / 2A \quad (5)$$

Here $y^*(n)$ is the complex conjugate of the output of the filter $y(n)$. To evaluate (5) for the Type 1 error, $y(n)$ needs to be expressed in terms of $\mathbf{f}(n)$. A complex equalizer will be assumed in the derivation.

$$y(n) = \sum_{l=0}^{N-1} f^*(n,l)x(n-l) = \mathbf{f}^H(n)\mathbf{x}(n) \quad (6a)$$

where

$$\mathbf{x}(n) = [x(n) \ x(n-1) \ x(n-2) \ \dots \ x(n-N+1)]^T \quad (6b)$$

is the input vector to the equalizer (adaptive filter) at instant n and $\mathbf{f}^H(n)$ is the complex conjugate transpose of the filter coefficient vector (tap weight vector) $\mathbf{f}(n)$. N is the filter length. Now, the expression for the instantaneous gradient becomes

$$\begin{aligned}
 \hat{\nabla}_{\mathbf{f}} J &= 2e(n) \frac{\partial e(n)}{\partial \mathbf{f}(n)} \\
 &= \frac{2e(n)}{2A} \frac{\partial}{\partial \mathbf{f}(n)} \{y(n)y^*(n) - A^2\} \\
 &= \frac{e(n)}{A} \frac{\partial}{\partial \mathbf{f}(n)} \{ \mathbf{f}^H(n) \mathbf{x}(n) \mathbf{x}^H(n) \mathbf{f}(n) \} \\
 &= 2e(n) \mathbf{x}(n) y^*(n) / A
 \end{aligned} \tag{7}$$

Substituting for $\hat{\nabla}_{\mathbf{f}} J$ in (3), the weight update equation for the Type 1 error can be written as:

$$\mathbf{f}(n+1) = \mathbf{f}(n) - \mu e(n) y^*(n) \mathbf{x}(n) \tag{8}$$

In (8) all the constants are lumped together into the single constant μ , the step size of the adaptation process. The same will be done for the other types of error equation.

3.2 Type 2 Error

$$e(n) = |y(n)| - A = \sqrt{y(n)y^*(n)} - A \tag{9}$$

The gradient of the cost function is now,

$$\begin{aligned}
 \hat{\nabla}_{\mathbf{f}} J &= 2e(n) \frac{\partial e(n)}{\partial \mathbf{f}(n)} \\
 &= 2e(n) \frac{\partial}{\partial \mathbf{f}(n)} \left\{ \sqrt{y(n)y^*(n)} - A^2 \right\} \\
 &= 2e(n) \frac{2y^*(n) \mathbf{x}(n)}{2\sqrt{y(n)y^*(n)}} = 2e(n) \frac{y^*(n) \mathbf{x}(n)}{|y(n)|}
 \end{aligned} \tag{10}$$

The weight update equation obtained from (3) now becomes:

$$\mathbf{f}(n+1) = \mathbf{f}(n) - \mu e(n) \frac{y^*(n) \mathbf{x}(n)}{|y(n)|} \tag{11}$$

3.3 Type 3 Error

$$e(n) = A \ln(|y(n)|/A) = \frac{A}{2} \ln(|y(n)|^2/A^2) \tag{12}$$

The gradient of the cost function is in this case,

$$\begin{aligned}
 \hat{\nabla}_{\mathbf{f}} J &= 2e(n) \frac{\partial e(n)}{\partial \mathbf{f}(n)} = \frac{Ae(n)}{2} \frac{\partial}{\partial \mathbf{f}(n)} \left\{ \ln(|y(n)|^2/A^2) \right\} \\
 &= \frac{Ae(n)}{|y(n)|^2} \frac{\partial}{\partial \mathbf{f}(n)} \left\{ |y(n)|^2 \right\} = Ae(n) \frac{y^*(n) \mathbf{x}(n)}{|y(n)|^2}
 \end{aligned} \tag{13}$$

The weight update equation, from (3), becomes,

$$\begin{aligned}
 \mathbf{f}(n+1) &= \mathbf{f}(n) - \mu e(n) \frac{y^*(n) \mathbf{x}(n)}{|y(n)|^2} \\
 &= \mathbf{f}(n) - \mu e(n) \frac{\mathbf{x}(n)}{y(n)}
 \end{aligned} \tag{14}$$

The weight update equations derived in this section are used to update the coefficients of the equalizer in our subsequent simulations.

4. CONVERGENCE RATE - SIMULATION

A Fractionally Spaced Equalizer (FSE) with an over-sampling factor of 3 ($M = 3$) was used in the simulations. For simplicity, BPSK modulation with rectangular pulse shaping was used. The block diagram of the simulated system is shown in Figure 3.

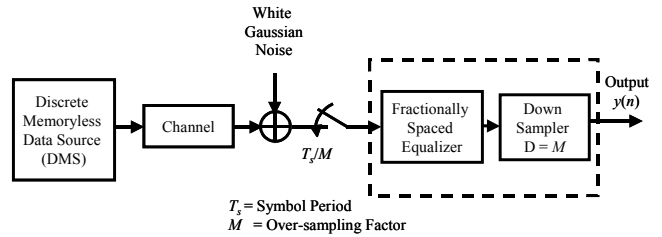


Figure 3: System Used for Error Signal Analysis.

The *fractionally spaced equalizer* (FSE) of Figure 3 is an adaptive equalizer whose coefficients are updated every M th output sample. This is because the FSE requires M samples of input to make a decision for a single output. For this reason, in Figure 3, the FSE and the down sampler are shown in the same block. All simulations were performed with external noise and the *signal-to-noise ratio* (SNR) used in simulation was 40 dB. To observe the effect of different error equations on the convergence rate of the equalizer, a symmetric channel [3] with the following impulse response is used in the simulation.

$$h(n) = \begin{cases} \frac{1}{2} \left[1 + \cos\left(\frac{2\pi}{W}(n-1)\right) \right], & n = 0,1,2 \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

The parameter W controls the amount of amplitude distortion produced by the channel, with the distortion increasing with W . The simulation uses $W = 3.1$. Also, $h(n)$ was normalized to have unit energy so that,

$$\sum_n h^2(n) = 1 \tag{16}$$

Now the FSE can be decomposed into multiple channels, where the number of sub-channels is the same as the over-sampling factor M . To have a perfect equalizer, the sub-channels (baud spaced) of all of the

paths should not have common zeros [4]. For this purpose, an extra zero was inserted (at time index $k=6$ in fractional space)¹ while converting the baud space channel to the fractional space channel. This extra zero makes the channel length equal to 10 in fractional space. The channel impulse responses, both in baud space and in fractional space, are shown in Figure 4.

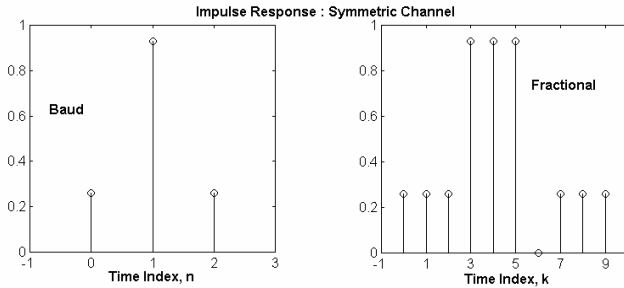


Figure 4: Impulse Response of Simulated Channel.

If the length of the channel (fractional space) and the equalizer are respectively ML and MN , then for perfect source recovery, N should satisfy the following condition [4]:

$$N \geq \frac{L-1}{M-1} \quad (17)$$

Therefore, for a channel of length 10 (in fractional space), and an over-sampling factor of 3, the length of the FSE should be at least 6, i.e., $MN \geq 6$. The length of the equalizer (FSE) used in the simulation was 12. The results of the simulation are shown in Figure 5.

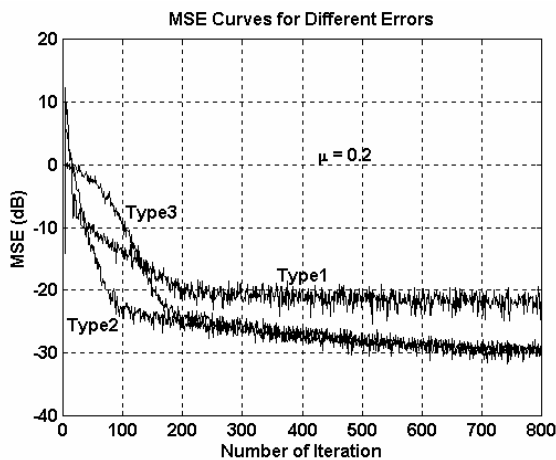


Figure 5: MSE Curves for Different Error Equations.

To generate the MSE curves in Figure 5, $\mathbf{x}(n)$ in weight update equations (8), (11) and (14) was replaced by $\mathbf{x}(n)/|\mathbf{x}(n)|^2$. An adaptation step size of 0.2 was used for

these simulations. An ensemble average of fifty independent realizations was used to generate each of the curves in Figure 5. From Figure 5 it is clear that, at the beginning of the adaptation process, the convergence rate is much faster for the equalizer with a Type 1 error equation. But after relatively few samples, the convergence rate with Type 1 becomes slower than for the equalizer with a Type 2 error. The reason for this is the non-symmetric behavior of a Type 1 equation around zero error. Due to its non-symmetric error behavior, the equalizer with Type 1 error provides worse steady state performance. We found that using a small enough step size effectively eliminates the effect of the non-symmetric error on the steady state performance, at the expense of a reduction in convergence rate. The Type 2 and Type 3 errors exhibit similar concave symmetry around zero error. We observe that the steady state MSE levels for these two types are almost the same. Equalization with a Type 3 error yields a slower convergence rate at the beginning, as anticipated in Section 2.

The same types of experiment were performed with different types of channel. All these experiments show effects of the different errors on the convergence rate of the equalizer analogous to those shown in Figure 5. Therefore we conclude that the Type 2 error equation provides the optimal solution for the blind equalizer since it yields a faster convergence rate and a lower steady state error, for a specific external noise, than a Type 1 error. The problem with a Type 2 error (and also with a Type 3 error) is that, for some channels, the estimate of the equalizer may be biased. To render the solution unbiased some kind of constraint needs to be used while minimizing the cost function.

5. EFFECT OF INITIALIZATION - SIMULATION

In a blind adaptation scheme, the initialization of the equalizer coefficients determines whether the equalizer will converge to a global or a local minimum. To show the effect of the initialization we will simulate a simple communication system, such as that shown in Figure 6. This choice facilitates comparison with earlier results [5].

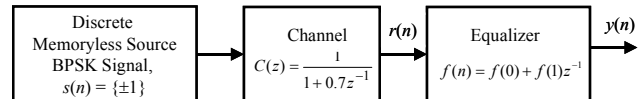


Figure 6: Simple Communication System with Adaptive Equalizer.

A baud spaced FIR equalizer of length 2 will be used to nullify the effect of the channel on the received signal. To converge to the correct minimum point, the equalizer

¹ Time index n is used in baud-space and k is used in fractional space.

coefficient vector should have the value of $\mathbf{f} = [1 \ 0.7]$. If a differential encoding scheme is used, $\mathbf{f} = [-1 \ -0.7]$ would also be an acceptable solution. Therefore, in the equalizer space (\mathbf{f} -space), the positions of the acceptable or global minima are $\pm[1 \ 0.7]$. If the adaptation process causes the equalizer to converge to other points in \mathbf{f} -space, those points will be considered to be local minima.

The effect of the initialization on the convergence behavior will be observed by initializing the equalizer coefficients at 20 different points in \mathbf{f} -space. These 20 points are selected such that they all are equally spaced and lying on a circle of radius 2. The same input sequence of length 4000 is used with all of the different initializations. For these simulations an adaptation step size of 0.005 is used in the original update equations (8), (11) and (14). The simulation results for different types of error are shown below.

Figures 7 and 8 show that with the Type 1 and Type 2 errors, convergence to global minima depends strongly on the initialization.

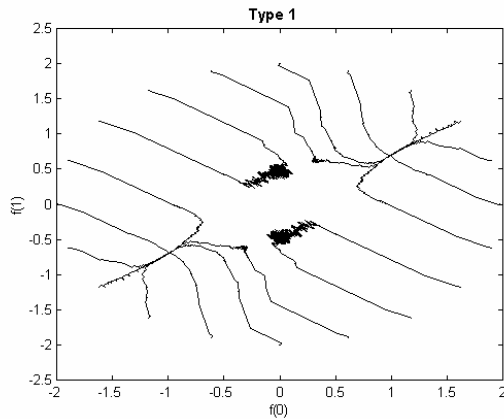


Figure 7: Effect of Initialization for Type 1 Error.

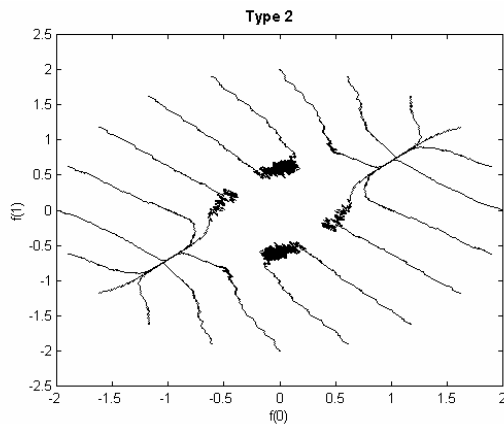


Figure 8: Effect of Initialization for Type 2 Error.

In Figure 9 the circles 'o' indicate the initial points and the asterisks '*' indicate the instantaneous value of the equalizer coefficients after 4000 iterations.

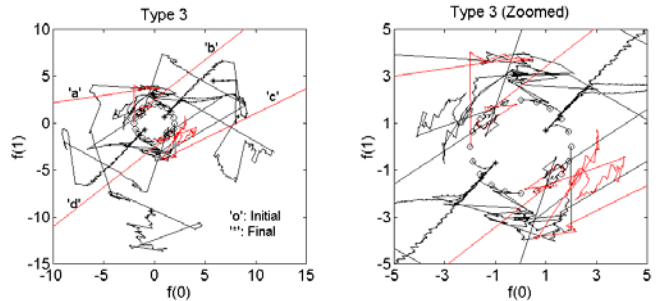


Figure 9: Effect of Initialization for Type 3 Error.

From Figure 9 it is clear that the Type 3 error exhibits a completely different property than the other error types. With a Type 3 error and a differentially encoded source, the equalizer will converge only to the global minima, irrespective of the initialization of the equalizer. However, the Type 3 simulation results also show problems with the convergence of the equalizer (cases a, b, c, and d in Figure 9). From (1) it is clear that if for any reason the output $y(n)$ becomes negligible, this will render the Type 3 error $e(n)$ very large ($e(n) = -\infty$ for $y(n) = 0$). In (14), the weight update equation, $y(n)$ appears in the denominator. Therefore, if $y(n)$ becomes too small, the equalizer coefficient vector will be modified by a large correction. As a consequence of these two facts, the equalizer will start to diverge. To avoid divergence of the equalizer, we now modify the Type 3 error equation as follows:

1. Use a hard-limiter at the output of the equalizer, which will not allow the output amplitude of the equalizer to go below a certain level. In the simulation this level was set to 0.01.
2. The error equation will be modified to:

$$e(n) = \begin{cases} \ln\left(\frac{|y(n)|}{A}\right), & \text{for } |y(n)| \geq A \\ -\ln\left(\frac{2A - |y(n)|}{A}\right), & \text{for } |y(n)| < A \end{cases} \quad (18)$$

The effect of initialization with the modified Type 3 error equation is shown in Figure 10. The same simulation parameters were used to generate the results in Figure 10 as were used for generating Figure 9. The same weight update equation, i.e. (14), was used with the modified Type 3 error equation.

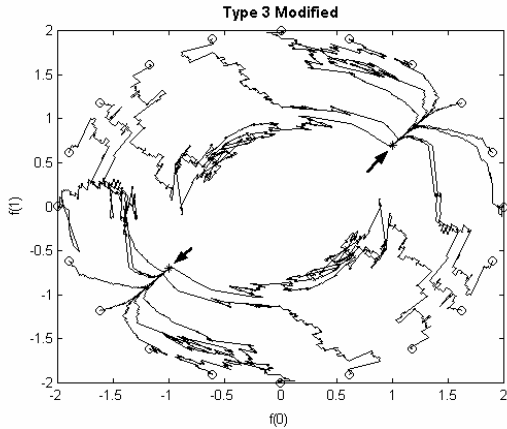


Figure 10: Initialization With Modified Type 3 Error.

It is evident from Figure 10 that, with the modifications, the logarithmic error equation leads the equalizer to converge to a global minimum (indicated by the arrows in the figure) and thereby eliminates the dependency of the convergence behavior on the initialization. Figure 11 shows the contour plot of the cost surface with the Modified Type 3 error, and confirms the absence of local minima in the cost surface of CMA with modified logarithmic error.

6. CONCLUSION

Based on the simulation results we conclude that CMA with an affine error equation exhibits better convergence rate than with either an exponential or a logarithmic error equation. A logarithmic error equation yields better results than the other two error types in the sense of converging to the global minima.

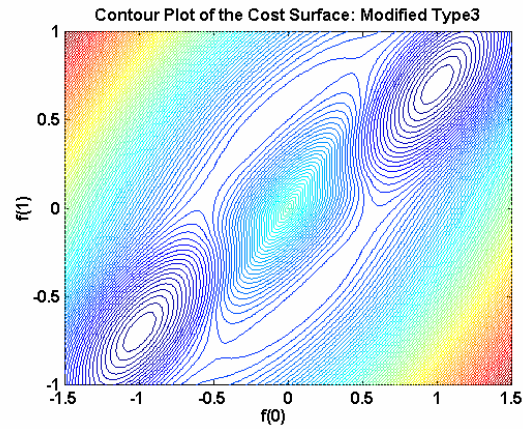


Figure 11: Cost-surface for Modified Type 3 Error.

REFERENCES

- [1] John R. Treichler, C. Richard Johnson, Jr., Michael G. Larimore, *Theory and Design of Adaptive Filters*, Wiley-Interscience, NY, 1987, pp. 157-170.
- [2] Scott C. Douglas, Teresa H.-Y. Meng, "Stochastic Gradient Adaptation Under General Error Criteria," *IEEE Transactions on Signal Processing*, vol. 42, no. 6, June 1994.
- [3] Simon Haykin, *Adaptive Filter Theory*, 2nd ed., Prentice Hall, NJ, 1991.
- [4] Ye Li and Zhi Ding, "Global Convergence of Fractionally Spaced Godard (CMA) Adaptive Equalizers," *IEEE Transaction on Signal Processing*, Vol. 44, no. 4, pp. 818-826, April 1996.
- [5] C. Richard Johnson, Jr., "Admissibility in Blind Adaptive Channel Equalization," *IEEE Trans. on Control Systems*, January 1991, pp. 3-15.