# THE USE OF EVOLUTIONARY OPTIMISATION IN CHANNEL EQUALISATION

*F. Sweeney, P. Power, C.F.N. Cowan*
Dept. of Electrical and Electronic Engineering,
The Queens University of Belfast,
Ashby Building, Stranmillis Road,
Belfast BT7 5AH, UK
Tel: 01232 274049;
e-mail: f.sweeney@ee.qub.ac.uk

**ABSTRACT**

This paper outlines the use of an Evolutionary Algorithm (EA) to perform the Equalisation of a non minimum phase channel. Conventional techniques utilising first and second order approximations of the error surface, have been demonstrated to be ineffective in achieving an optimal solution in continuous simulations, and have proved incapable of dealing with the more difficult non minumum phase problems. Using an EA, this paper will show how a consistent, near optimal, solution can be achieved.

## 1 INTRODUCTION

EAs offer a global optimisation technique that is ideal for the multimodal, nonlinear optimisation of Multi-Layer Perceptron (MLP) Neural Networks to solve the channel equalisation problem. In recent years there has been considerable research into the problem of channel equalisation, and in particular the application of non linear solutions to this problem *[1,2,3]*.

However, this research has been considerably hampered by the lack of adaquate algorithms to optimise these non linear structures. Current procedures utilise a gradient descent approach (Back Propagation) to the optimisation problem, using first or higher order approximations of the error surface to aid in the estimation of a better, more optimal solution *[3]*. Unfortunately, to implement these techniques there needs to be some a priori assumptions as to the nature of the error surface. These assumptions can be summarised as:

- a differentiable error surface.
- a monomodal surface with a zero terminating gradient.

For an MLP structure, these assumptions are invalid. The error surface can be easily shown to be multimodal *[7]*, with very flat *plateaux* and steep *cliff* edges (Figure 1). It is primarily because of this topology that gradient descent has difficulty in the optimisation of MLP's; the slow learning in the plateau areas are followed by steep slopes that drive gradient descent into instability. Another global optimisation technique is required, one that is better suited to the difficulties in the optimisation surface.

An (EA) offers a global search technique that has had some success with difficult optimisation problems. EA's utilise Darwinian *survival of the fittest* ideology to find an optimal solution. They use a population of individuals, called a *chromosome*, and by a process of selection and breeding of each individual, called *genes*, a best gene will emerge. Individuals are selected by merit or *fitness*, with the more meritorious individuals having a greater chance of propagation to latter generations. To successfully immitate a system of organisms, there is also a mutation operation occuring on selected genes within the chromosome.

To summarise, this paper will describe how an EA works, as well as highlight modifications to MLP structures that have produced improved results when compared to the standard MLP implementation. A comparison of the EA with a standard Back Propagation (BP) algorithm will show how the EA outperforms BP by using a Bit Error Rate (BER) performance comparison over an ensemble of simulations using the two techniques.

## 2 DESCRIPTION OF EA

An EA is a stochastic search technique untilising Darwinian criteria to improve the fitness of a population of genes *[4,5]*. To find an optimal configuration, the EA relies on three basic operations: *survival*, *crossover* and *mutation*.

However, before the optimisation can proceed it is necessary to encode the MLP into a format suitable for optimisation. This is achieved by initialising the population with a structure that is suitable for the optimisation *[5,6]*. Previous work has shown that it is possible to initialise a network to fit a certain topology, that is specific to the problem, without significantly decreasing the generality of the MLP in the optimisation. Once this is done the optimisation can proceed to find the best configuration *[6]*.

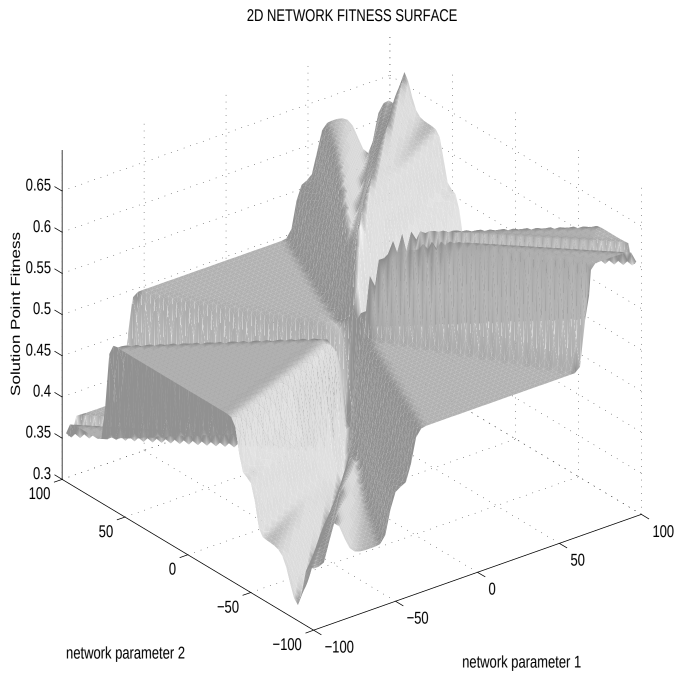


Figure 1: Error Surface of an MLP, highlighting the difficulty in optimisation by conventional gradient descent techniques

### 2.1 Encoding of the EA

A specific encoding is used to ensure the EA can efficiently optimise the MLP. If we assume a network made up of hardlimiting devices, we can successfully dychotomise the MLP into two distinct processing features. If the input layer is considered, the network input will be a time delayed sampled channel consisting of channel output and noise; this is, in effect, a sampled analogue channel. Now, since the processing nodes utilise the hard limiting activation function, the ouput of the first layer will be a binary vector. Therefore, the function of the MLP can be described thus: the first layer is primarily an analogue to digital converter, whilst the latter layers perform a Boolean classification of this data.

To summarise, by assuming whether the channel is minimum phase or not, the outer layers can be used to perform a static logic mapping, whilst the input layer perceptrons try to achieve the best partition of the input space *[6]*.

### 2.2 Initialisation

To initialise the EA, each member of the population (i.e. each MLP) is instantiated with a specific topology that is similar to the optimum *shape*. It has been shown that using a specific topology, an increased performance from the EA can be achieved without reducing the learnibility of the MLP. The BP optimisation was initialised using the same strategy to ensure a fair comparison.

These MLP configurations are created by initialising the gradients, and biases, of the the perceptrons to a parameter range that has a topology similar to the solution. The EA will then optimise the decision boundaries to achieve the best configuration. The biases of the MLP were intialised to random values between $\pm 2$ with the weights initialised to values between $\pm 5$. To create a chromosome, the weights and biases of the MLP are represented as an N dimensional vector, i.e. each weight or bias represents a dimension.

### 2.3 Survival

To achieve an optimal configuration of weights, it is necessary to remove the poorer performing individuals. To do so, it is imperative to guage the effectiveness of individuals within a population. This can be achieved by showing each member a training data set, and using this to calculate the errors associated with each individual. They can then be ranked according to fitness *[4]*.

A random selection criteria, weighted by the fitness of each individual, is used to select which networks will be allowed to continue to the next generation. This process results in the fittest individual having a greater chance to crossover their individual genetic information.

### 2.4 Crossover

Once the better genes have been selected, a method of transferral of information needs to be implemented; to do so requires pairs of parents within the population to swap their genetic information. A single point crossover strategy was implemented within the EA, whereby a point of crossover was randomly selected between the two *parent genes*, resulting in two offspring genes replacing both parents *[5,6]*.

There is a further operation that actually swaps the information of the the two subgenes of the parents, with this ensuring a greater search potential of the population *[5,6]*.

### 2.5 Mutation

Mutation is the final parameter that adds a degree of randomness to the search technique. Without mutation, the search would be constrained to be within the boundaries of the parameters set at initialisation. Mutation can therefore be viewed as adding small perturbations to the system, and therefore ensuring a greater amount of the error space is explored by the EA.

The mutation rate was set to 10% for each individual, the mutation rate for each weight was set to 35%, which ensures that when a mutation takes place the individual is mutated to a significant extent. These values have been chosen quite arbitrarily, and there is scope for future investigation to determine how effective the current configuration is.

## 3 SIMULATION PROCEDURE

To investigate the feasibility of the EA, two different non minimum phase channels have been chosen. For the first simulations the channel model was Equation 1, whilst for the second set Equation 2 was used. In the first set of simulations a channel was equalised by an MLP trained by BP and an EA in order to compare the performance of the different algorithms. Ten simulations were performed for each SNR. The resultant networks were then subjected to a bit error rate test to determine how effective each of the training techniques were.

The second set of simulations used a slightly different network configuration. The input and hidden layers were made up with hard limiting perceptrons, whilst the output layer was made up of a softer hyperbolic tangent activation function. The aim of this configuration was to slow down the convergence to enable the MLP to converge to a more optimal value. Ten simulations were run with a varying gradient parameter of 0.1, 0.5 and 1.0; these were compared with the optimal achieveable performance *[1]*.

The aim of the simulations was to show that a softening of the output layer would improve the overall BER performance, by reducing the likelihood of a suboptimal configuration. The BER performance for each of these simulations was ensemble averaged to achieve an expected BER performance for an MLP of a specific configuration. Below are the z transfroms of the two channels.

$$H(z) = 0.2362 + 0.8636z^{-1} + 0.2362z^{-2} \tag{1}$$

$$H(z) = 0.3482 + 0.8704z^{-1} + 0.3482z^{-2} \tag{2}$$

## 4 RESULTS

The results indicate that the EA provides a more effective means of training an MLP to perform the equalisation of a non minimum phase channel. Figure 2 shows that the EA would train the MLP to a more optimal solution more often than BP. The plot also shows that the performance is slightly less than optimal. On average the variance of the BER performance of MLPs trained with an EA was much less than those trained with BP.

Another interesting point was that, whilst the EA and BP could train a MLP to an optimal configuration, the complex nature of the error surface posed little difficulty for the EA, whilst BP would suffer stability problems and not converge to a satisfactory solution. There was also a difficulty in determining a satifactory termination criteria for BP, whilst the EA would converge satisfactorily within 15 generations (the limit used to calculate the optimal network using an EA).

Figure 3 shows the effect of using a dual structure network, i.e. a network with hardlimiting activation functions in the input and hidden layers and a hyperbolic tangent function in the output layer. The reasons for this configuration is to enhance the noise characteristics of the EA. If a network is made up of hardlimiting perceptrons, it tends to be impervious to the noise of

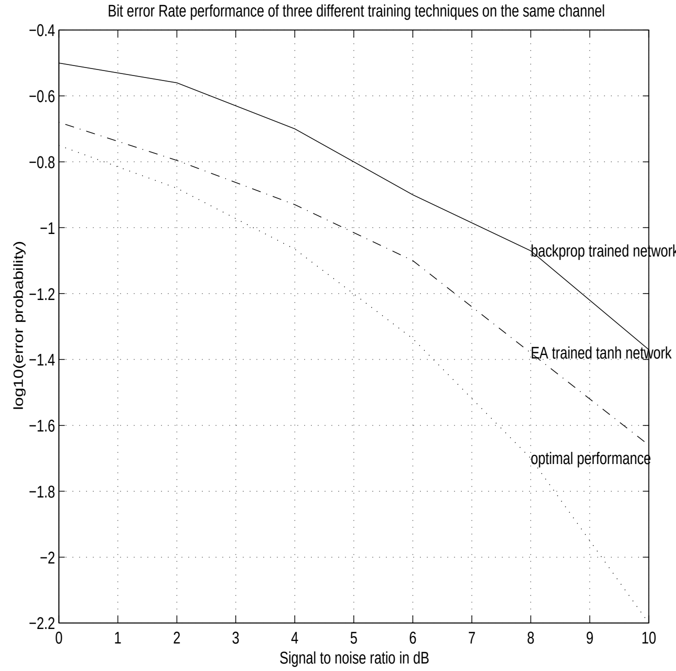


Figure 2: showing the performance of the EA compared to Back Propagation

the training sequence, and therefore can cause the MLP network to converge prematurely to a suboptimal configuration. This can be alleviated by softening the ouput perceptron, to ensure there is always an error with the output to add a degree of uncertainty.

## 5 DISCUSSION

These results when coupled with previous research show that EAs offer an effective alternative training technique for MLP's to perform channel equalisation. EA's have been useful in training MLPs of different structure and do in fact outperform conventional gradient descent techniques. This has been verified in two separate cases, the equalisation of difficult non minimum phase channels using MLP structures that were thought of as being incapable of learning a specific boundary to the degree that was thought necessary *[3]*, and the BER analysis with time delayed channels to compare with conventional gradient descent techniques. In all of these analyses the EA outperformed gradient descent, primarily because of the inability of gradent descent to deal with the difficult error surface shown in Figure 1.

EAs do have a number of flaws that results in a conflict of interest to the user, these are:

- A fast convergence that can result in the premature convergence to a suboptimal solution.
- By slowing down the convergence a better more optimal solution can be obtained.

The convergence of the EA can be hastend up in two

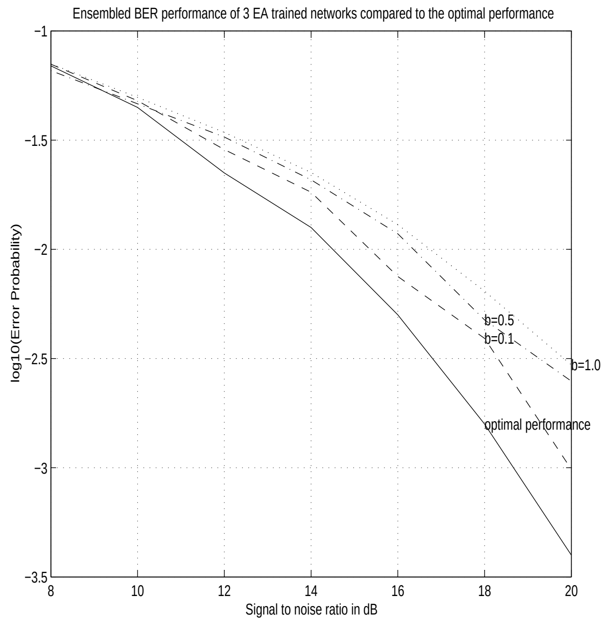


Figure 3: showing the performance of the EA compared to Back Propagation

ways: by increasing the selectivity of the fitness function, and by making the output layer a hardlimiter activation function for the training cycle. The first operation speeds up convergence by ensuring that the current fitter individuals have a greater chance of propagating their genetic information into subsequent generations, whilst the second operation tends to ignore the presence of noise in the optimisation and in doing so cause the premature convergence of the EA.

The softening of the output (i.e. by making it a hyperbolic tangent) function is beneficial primarily because it incorporates noise information into the optimisation process whilst maximising the learnability of the current structure through the use of hardlimiting activation functions in the hidden layers *[8,9]*. After the training cycle the softlimiting activation function is replaced by the conventional hardlimiter; this ensures that the ouput will be either $\pm 1$.

## 6 CONCLUSIONS

Evolutionary Algorithms offer a global search technique that is highly effective in training nonlinear structures (MLPs) to perform the equalisation of difficult nonminimum phase channels. It has been shown to outperform a standard gradient descent technique (BP) over a range of simulations on two separate channels. However, a question remains unanswered; why does an EA outperform BP? The simple answer is due to the multimodal nature of the optimisation surface, which means that BP (an essentially monomodal optimisation strategy) has difficulty in locating an optimal point. Due to the global nature of the EA search, a better view of the optimisation surface can be created, with this aiding the optimisation. Another difficuly with the error surface is its range of gradients, this has been shown by Figure 1.

Previous research has also shown that there has been considerable difficulty in the estimation of poles close to the unit circle by gradient optimisation techniques [10]. There is therefore a relationship between the error surface and the difficulty in the equalisation problem, this further discounts the effectiveness of conventional gradient descent based optimisation operations.

## 7 FUTURE WORK

Future work will involve the development of hybrid techniques, coupled with more precise cross over and mutation operations. The dimension of the input also needs to be increased to enhance the noise performance of the system to more acceptable levels; a more generalised initialisation procedure needs to be implemented for this purpose. Finally, the EA needs to be modified for the optimisation of a non static MLP structure, this will more efficiently utilise the EA's processing power.

## 8 REFERENCES


1. S. Siu, "Non-Linear Adaptive Equalisation based on Multi-layer Perceptron Architecture", PhD Thesis, Faculty of Science, University of Edinburgh, 1990.
2. S. Siu, G.J. Gibson, C.F.N. Cowan, "Decision Feedback Equalizer using Neural Network Structures and Performance comparison with Standard Architecture".
3. S. Haykin, "Neural Networks, A Comprehensive Foundation".
4. Goldberg D.E., "Genetic Algorithms in Search Optimisation and Machine Learning", Addison Wesley.
5. F. Sweeney, C.F.N. Cowan, "The Use of Evolutionary Algorithms in Channel Equalisation", submitted for ICASSP'98, Seattle.
6. F. Sweeney, C.F.N. Cowan, "The Use of Evolutionary Techniques in Channel Equalisation", CSDSP'98, pp36-40, Sheffield Hallam University.
7. P. Power, C.F.N. Cowan, "Genetic Neural Networks", Internal Paper, Department of Electrical and Electronic Engineering, The Queens University of Belfast, 1997.
8. G.J. Gibson, C.F.N. Cowan, "On the Decision Regions of Multilayer Perceptrons", Proceedings of the IEEE, vol.78, pp 258-262.
9. J.Connolly, "Equalisation using Neural Network Structures trained by Genetic Algorithms". M.Eng Thesis, Dept. Elecrical and Electronic Engineering, The Queens University of Belfast.
10. Quiang Ma, "The Application of Genetic Algorithms to the Adaption of IIR Filters", PhD Thesis, Department of Electrical and Electronic Engineering, Loughborough University of Technology, 1995.


## 9 ACKNOWLEDGEMENT

C.F.N. Cowan holds a professorship sponsored by Nortel and The Royal Academy of Engineering.