Journal of the IETE
Vol 40, Nos 2&3, March-June, pp 113-118

A Neural Net Based Scheme for Narrowband Signal Extraction

This paper proposes an alternative method for extraction of narrow band signal contaminated with varying white noise power. A three layred feedforward network with one hidden layer is suggested in this paper. The internal parameters are updated by employing the widely used backpropagation algorithm and the statistical Cauchy's algorithm. The network uses the tangent hyperbolic function to provide the desired nonlinearity to the scheme. During learning the net is exposed to signals with high SNR ranging from +40 dB to +20 dB. However, after learning quite interesting results are obtained for very low SNR of the input signal even up to 0 dB. The training samples for fundamental, 3rd harmonic and 5th harmonic are chosen to be different with a view to exploit the pattern classification feature of the net. Therefore, after termination of training the proposed scheme accomplishes the filtering action successfully. The phenomenon of network paralysis as encountered in BP algorithm is circumvented by opting for the statistical Cauchy's algorithm with Boltzmann's probability feature. The extraction capability of the proposed scheme is highlighted by providing inadequate noisy input samples to the net. Besides, a performance comparison is presented for both the algorithms. Results presented emphasizes the substitution of the scheme for adaptive filter for a particular zone of operation.

THE usual method of estimating a signal corrupted by additive noise is to pass it through a filter that tends to suppress the noise while leaving the signal relatively unchanged. The filters for the above purposes may be fixed or adaptive, where the later ones have the ability to adjust their own parameters automatically and their design requires little or no apriori knowledge of signal or noise characteristics. In the circumstances where adaptive noise cancellation is applicable [1] levels of noise rejection are attainable that would be difficult or impossible to achieve by direct filtering. A good number of adaptive algorithms have been reported in the literature over the past decades to update the internal parameters with a view to obtain high fidelity of the output signal in many complex environments. However, the major focus of research over the past few years is to develop robust algorithms for adaptive systems.

The explosive growth in the studies of the neural network motivated the researchers to find applications in the fields of speech, image recognition, signal processing [2-6] etc. In addition to the above research, a few research papers are reported exclusively in the areas of signal processing [8-11]. Also multitone detection is suggested by SS Rao [12]. J W Watterson [13] suggested the signal detection using multilayered neural network. In contrast this paper suggests a scheme for narrowband signal extraction. The net's internal parameters are updated by employing the backpropagation and statistical Cauchy's algorithm [14]. The tangent hyperbolic activation function is selected here to suit the signals considered. The network consists of input points equal to the number of input samples. The signals considered encompass the power frequency signal samples. The number of input and output nodes are selected as 16 which implies that the net is exposed to sixteen samples at a time. The signal is sampled at such a sampling frequency that at a time one cycle of waveform is considered. The same concept is extended to determine the number of output nodes. The number of hidden nodes are selected in a trial and error basis to accomplish effective generalization of the net. But in this scheme the number of input, hidden and output nodes are the same. The signal samples are slided so that the net is exposed to the same number of samples all the time with one new sample added and the oldest one is neglected from the set. The net is trained for fundamental, third harmonic and fifth harmonic components. Subsequently, the signal samples consisting of either of the above mentioned frequency components and noise are applied to the net. These samples when fed through the net layers is found to produce quite consistent signal samples at the output nodes. Appreciable results are also observed at the output nodes with the incomplete signal sample set. Surprisingly the scheme is able to extract the desired samples with 50% suppressed input data both with predominant dc component and appreciable harmonic strength. The network paralysis with the backpropagation algorithm demands to switch over to the statistical algorithm. A performance comparison is made between the two algorithms to highlight the convergence characteristics and local minima trapping effect. Also simulation study on faulty net is carried out and quite promising results are obtained. Computer simulation results are presented in this paper for a wide variety of cases.

Paper No 154-B; received 1991 January 2; revised 1991 October 20;

CONCEPT BEHIND THE PROPOSED SCHEME

The neural network used is a three layered network with input, hidden and output layer as shown in Fig 1. The hidden units are chosen to be the same as that of the input and output units. The sample sets are used one after another. In each adaptation a new sample is added and the oldest one is discarded thereby keeping the length of the sample set fixed. So the samples are slided and hence the network is trained for a particular signal. Similar concept is extended for the 3rd and 5th harmonic cases. The signals considered for the three cases,

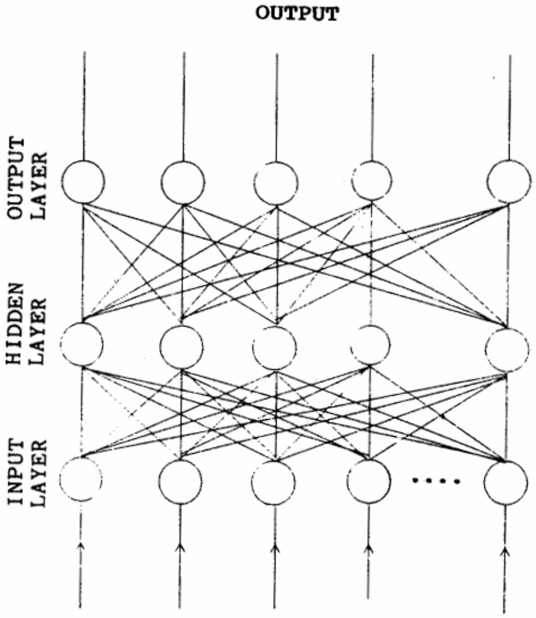

Fig 1 A typical three layered feedforward artificial neural network

ie, fundamental, 3rd and 5th harmonic are different. After the network is properly trained the signal samples with high noise power are applied to the input nodes. Since the net is trained properly, its generalization attribute helps to extract the desired component.

TRAINING ALGORITHM

The net trained to justify the signal extraction capability is shown in Fig 1. The subscripts *i, j* and *k* refer to the input, hidden and output layers respectively, and the subscripts *n, m* and *l* correspond to any unit in input, hidden and output layers. The net is trained with supervision by the required input and target vectors.

Backpropagation Algorithm

The weights of both the layers are initialized to small random values. The input vectors x_0, x_1, ..., x_{N-1} and the target vector d_0, d_1, ..., d_{N-1} are formed and applied to the net. Here, in this research, the input and the desired vector are of same length. At each time step the following equations are computed

$$\hat{y} = f\{\lambda x [w]\} \tag{1}$$

where

$\hat{y}$ = estimated output of any layer.

λ = weighting factor to constrain the activation function in the nonlinear zone.

$[w]$ = weighting matrix between any two layers.

and, $f(.) = A \tanh(.)$

where

A = maximum value of the activation function.

$$e = d - \hat{y} \tag{2}$$

$$E_p = (1/2) \sum_{k=1}^{1} (d_k - \hat{y}_k)^2 \tag{3}$$

where

e = error between actual and target.

E_p = objective function.

The weights of any layer are updated as per the equation given below

$$[w]_{ij}(t+1) = [w]_{ij}(t) + \eta\, \delta_j x_{-j} + \alpha\{[w]_{ij}(t) - [w]_{ij}(t-1)\} \tag{4}$$

where $[w]_{ij}(t)$ is the weight from the hidden node i or from an input to node j at time t, x_{-j} is either the output of node i or is an input, η is a convergence coefficient varied from 0.1 to 1.0, δ_j is an error term of node j and α is the momentum term which varies from 0.0 to 1.0 and it takes care of the effect of previous weight on the current weight. If node j is an output node then

$$\delta_j = (A - \hat{y}_j^2 / A)(d_j - \hat{y}_j) \tag{5}$$

where, d_j is the desired output of node j and y_j is the actual output. If node j is an internal hidden one, then

$$\delta_j = (A - x_j^2 / A) \sum_{k=1}^{1} \delta_k [w]_{jk} \tag{6}$$

where, k is over all the nodes to which j transmits the signal.

Cauchy's Algorithm

The following steps are adopted for training of the network.

Step 1: A variable T_0 that represents an initial artificial temperature is chosen. Usually T_0 is initialized to a large value. Also another variable t = artificial time analogous to a step, is initialized. Usually t is set to a small value. Here T_0 = 9898.0, t = 2.0. The differential weight computed depends upon T_0 and also the convergence time is large, hence T_0 is selected to be a large arbitrary number, *ie*, 9898.0.

Step 2: A set of input is applied to the network, the corresponding outputs and the objective functions are evaluated.

Step 3: The weights are changed by a random value which is given as

$$x_c = \rho\{T(t) \tan [P(x)]\} \tag{7}$$

where

$$T(t) = \frac{T_0}{1+t}$$

ρ = the learning rate coefficient.

x_c = the weight change.

$P(x)$ = a random number selected from a uniform distribution over the open interval $-\pi/2$ to $+\pi/2$.

Step 4: If the objective function is improved (*ie*, reduced), retain the weight change.

$$E_p = 1/2 \sum_{j=1}^{1} (d_j - \hat{y}_j)^2 \tag{8}$$

Step 5: If the weight change results in an increase in the objective function, calculate the probability of accepting that change from the Boltzmann's distribution as follows

$$P(c) = \exp(-c/kT) \tag{9}$$

where $P(c)$ is the probability of a change of c in the objective function, k a constant analogous to Boltzmann's constant that must be chosen. Here, $k = 10^{-17}$ is selected and T is the artificial temperature.

The above steps are repeated until the objective function is minimized.

SIMULATION

In this research the network is trained independently for each cases namely, fundamental component of 50 Hz, the corresponding 3rd and 5th harmonic, etc. The signal model considered is

$$y(t) = \sum_{m=1,3,5\ldots} \sin(m\omega t) + V(t) \tag{10}$$

where m = 1, 3, 5, 7,....., n and $V(t)$ = Zero mean white noise. The following training sample sets are generated

* Fundamental signal (50 Hz) + noise ($-$ 40 dB and $-$ 20 dB noise power down the signal power).

* Fundamental + 3rd harmonic + noise ($-$ 40 dB and $-$ 20 dB noise power down the signal power).

* Fundamental + 3rd harmonic + 5th harmonic + noise ($-$ 40 dB and $-$20 dB noise power down the signal power).

The corresponding target components for the above three cases are the fundamental, 3rd harmonic and 5th harmonic respectively. The sampling frequency for each of the above mentioned cases is 800 Hz whereas the fundamental frequency is 50 Hz. One cycle of the signal is considered at a time. Hence, 16 samples form a sample set for training. The number of input points selected for the network is 16 and same number of hidden and output units are taken. In the next adaptation the net is exposed to a sample set by neglecting the first sample and adding a new sample to the set. However, the target samples are unaltered for a particular signal component. Thus the net is exposed to the sets for learning. Now, the first sample set with the corresponding target set is impressed to the net and the network parameters are updated. In a similar way the other two cases, *ie*, 3rd and 5th harmonics are dealt with. For each case the net is trained with two different noise powers, *ie*, SNR = 40 dB and 20 dB. In each case the total signal power is always unity. Now the noise power is varied by the necessary routine, thereby monitoring the required SNR. The different parameters in the backpropagation algorithm namely η, the slope of the activation function λ, maximum value of the activation function A and the momentum term α are varied to obtain the desired floor level of the objective function. In Cauchy's algorithm the parameters like learning coefficient ρ, λ and A are varied to achieve effective training. Since Cauchy's algorithm is based on the simulated annealing principle, T_0 represents that an initial temperature is chosen to be a high value, *ie*, 9898.0.

The number of adaptations required for proper training in case of statistical algorithm is more than that of the backpropagation algorithm. The different weights computed in each adaptation depends on the instantaneous value of the temperature. Hence, an arbitrarily high value of T is chosen.

RESULTS AND DISCUSSION

The learning curve for fundamental component employing backpropagation algorithm is depicted in Fig 2. The objective function attains a floor level of -100 dB after approximately 400 adaptations. The observed spikes in the Fig 2 correspond to the net's encounter to a new input pattern. The corresponding target vector is presented in Fig 3. The corrupted input signal and the net's extracted outputs are presented in Fig 4. The output resembles that of the target components, thus implies enhanced signal extraction capability. The net's potentiality for signal extraction is reinforced in Fig 5 that shows incomplete input signal and the corresponding output. Figure 6 and Fig 7 exhibit the convergence characteristics for third harmonic with backpropagation and the statistical algorithm. The objective function's floor level value in Fig 6 is around -10 dB, thereby, displaying the phenomenon of local minima trapping. But in Fig 7 the objective function's floor level is around -35 dB down and that reveals the alleviation of the phenomenon of network paralysis. The floor level of the objective function is poorer in case of 3rd harmonic than that of the fundamental because of its signal condition for training. The interfering signals, *ie*, fundamental and the noise components are treated as unwanted signals for the 3rd harmonic case. Hence, the high power unwanted signals to be suppressed are responsible for poor convergence characteristic.

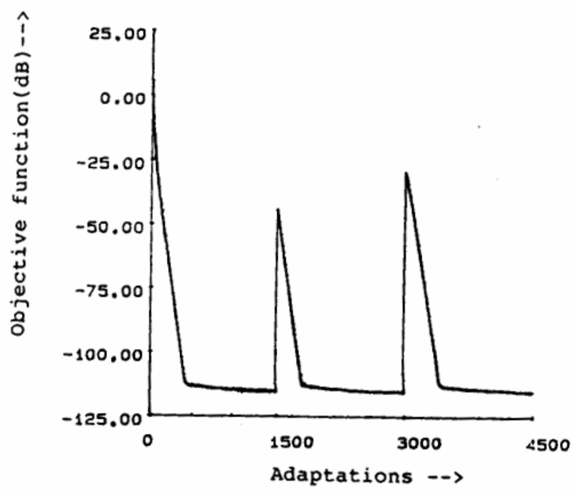

Fig 2 Learning curve for fundamental samples using backpropagation algorithm (three sets of training signals). $\eta = 0.45$, $\lambda = 0.0065$, $A = 3.0$, $f_c = 800$ Hz, $\alpha = 0.2$

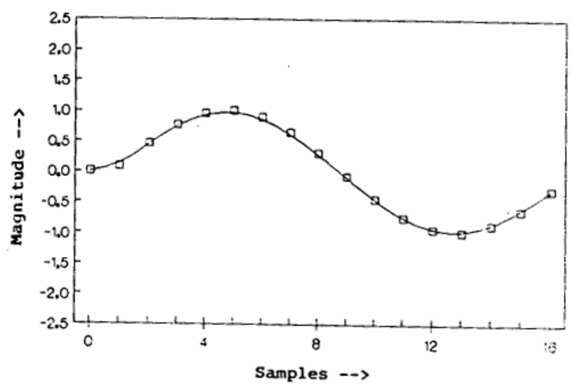

Fig 3 Target waveform for fundamental (50 Hz)

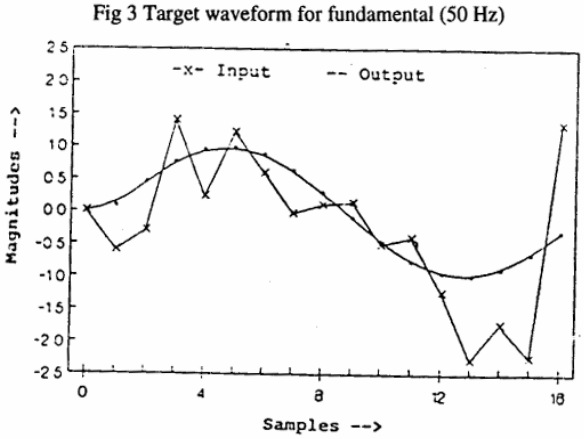

Fig 4 Input signal with SNR = 0 dB and the predicted output for fundamental

Figure 8 shows the target components for the third harmonic components. Figure 9 and Fig 10 exhibit the input signal with required SNR and the corresponding predicted ones. The predicted output with Cauchy's algorithm is more promising than that of the backpropagation algorithm. Figure 11 shows the convergence characteristics for the 5th harmonic component with backpropagation algorithm. The objective function attains a floor level of around – 40 dB due to the input signal condition. The adaptation of weights for the above case is faster than that of the fundamental one. The weights are also less sensitive to a new set of data as compared to the fundamental component. Figure 12 shows the desired signal for the 5th harmonic. Figure 13 present quite interesting output with incomplete and complete information respectively.

The above results establish the fact that the proposed scheme possesses better noise suppression capability than the conventional algorithms. The same effect is also observed for 3rd harmonic and 5th harmonic cases. However, it is constrained to operate successfully for narrow band signals. Once the net is trained, it is capable of extracting without invoking any algorithm. Hence, the computational burden as compared to adaptive algorithms which estimate the signal for each adaptation is less. Although Cauchy's algorithm takes 1000 adaptations to converge, still it is capable of producing

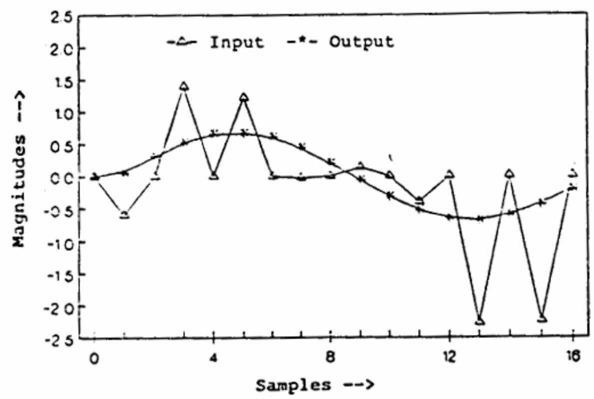

Fig 5 Input signal, SNR = 0 dB with even samples (*ie* 2, 4, …, 16) suppressed and the corresponding predicted output for fundamental

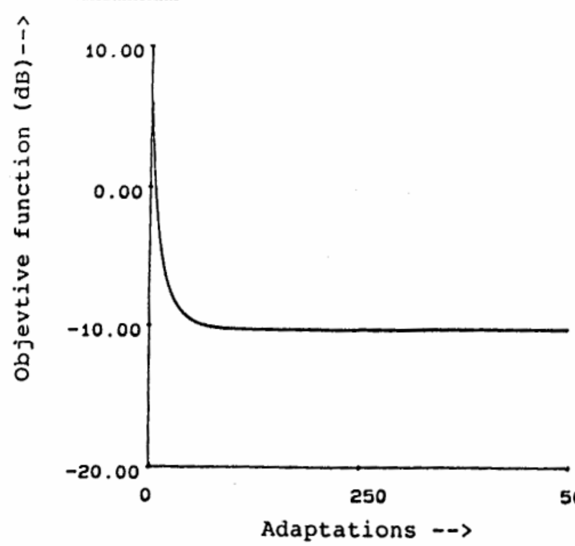

Fig 6 Learning curve for third harmonic (backpropagation algorithm) $\eta = 0.5$, $\lambda = 0.009$, $A = 2.4$, $\alpha = 0.4$

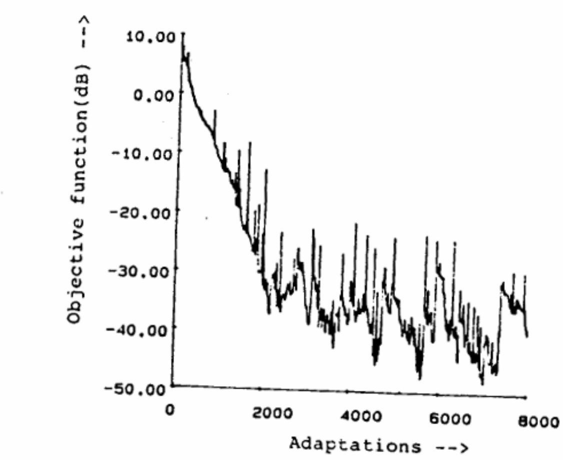

Fig 7 Learning curve for third harmonic using Cauchy's algorithm, $\rho = 0.01$, $\lambda = 0.07$, $A = 2.0$

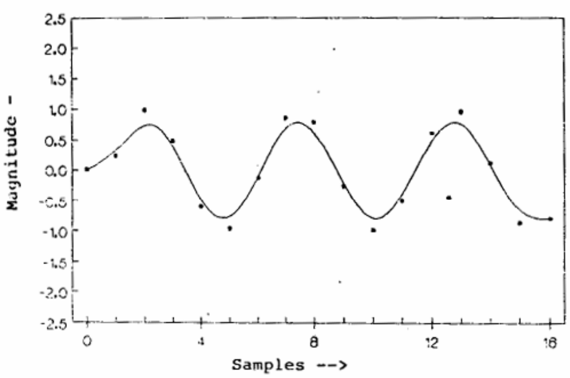

Fig 8 Target for third harmonic

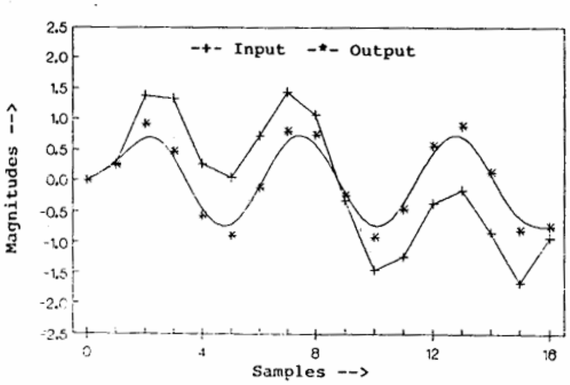

Fig 9 Input signal, SNR = 20 dB and output (backpropagation algorithm)

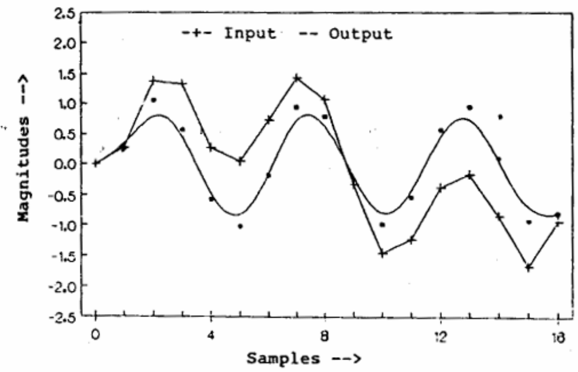

Fig 10 Input signal, SNR = 20 dB and output (Cauchy's algorithm)

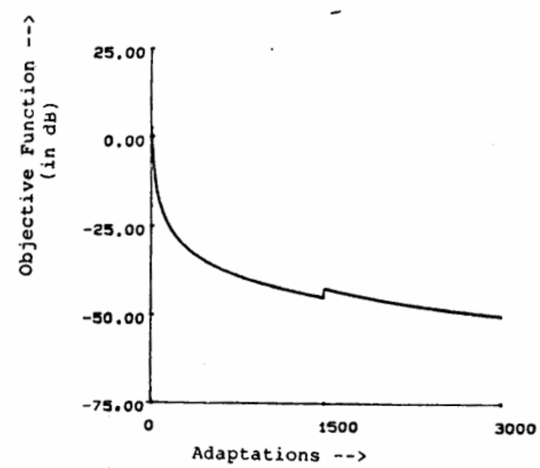

Fig 11 Learning curve for fifth harmonic (backpropagation algorithm), $\eta = 0.6$, $\lambda = 0.007$, $A = 3.0$, $\alpha = 0.3$

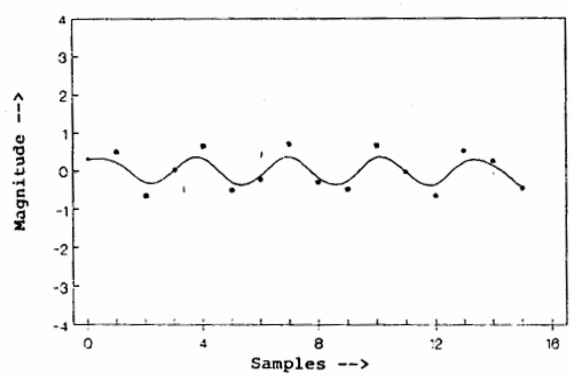

Fig 12 Target for fifth harmonic

appreciable results. This occurs because the differential weights are computed based upon random number distribution.

CONCLUSION

Exhaustive computer simulation results establish the validity for opting the proposed scheme over the existing digital filter. Simulation study shows the efficacy of the scheme for a particular zone of operation. The performance in extracting the signal components from inadequate corrupted data excel over the existing techniques. The local minima trapping effect of the objective function as encountered in backpropagation algorithm is circumvented by switching over to Cauchy's algorithm. The problem of the net's paralysis in Cauchy's algorithm is overcome by introducing the Boltzmann's probability distribution. The combined algorithm developed, overcome the local minima effect and the convergence is found

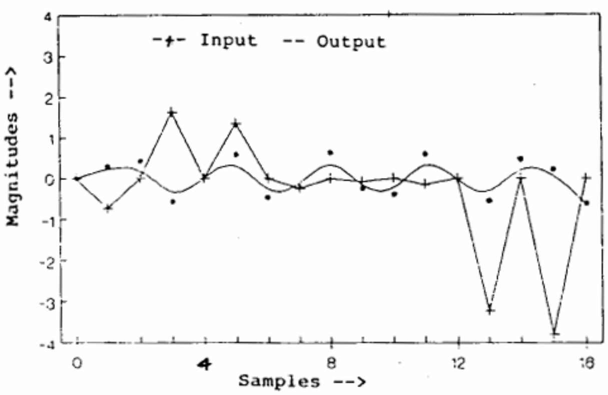

Fig 13 Input signal, SNR = 0 dB with even samples suppressed and the output

to be faster than that of the Backpropagation algorithm. The net is found to be fault tolerant and real time implementation of the generalized filter is in progress by employing both adaptive pattern recognition concept in conjunction with robust training algorithm.

ACKNOWLEDGEMENT

The authors acknowledge with thanks the funds available from the Ministry of Human Resources Development, Government of India for carrying out this work.

REFERENCES

1. B Widrow *et al*, Adaptive Noise Canceling Principles and Applications, *Proc IEEE*, vol 63, pp 1692-1716, December 1975.

2. R Lippman, An Introduction to Computing with Neural Nets, *IEEE ASSP Mag*, vol 4, no 2, pp 4-22, April 1987.

3. R Lippman, Pattern Classification using Neural Network, *IEEE Comm Mag*, pp 27-62, November 1989.

4. M R Azimi-Sadjadi *et al*, Neural Network using Fast Least Squares, *ICASSP*, April 1990, New Mexico, USA.

5. P Burrascano *et al*, A Learning Rule Eliminating Local Minima in Multilayer Perceptrons, *ICASSP*, April 3-6, 1990, New Mexico, USA.

6. P K Dash, P K Nanda & S Saha, Spectral Estimation using Artificial Neural Network, *Proc of Int Conf on Signals and Systems*, December 10-12, 1990, Tirupati, India.

7. C Wang, M Wickert & CH Wie , Three layer Neural Network for Spectral Estimation, *Proc of ICASSP*, April 3-6 1990, New Mexico.

8. Y Yin, Adaptive Neural Filters, *Proc of IEEE Workshop on Neural Networks for Signal Processing* held on 29th Sept-2nd October 1991 at Nassau Inn, Princeton, NJ, USA.

9. N Hoffman, A Neural Architecture for Non-linear Adaptive Filtering of Time Series, *Proc of IEEE Workshop on Neural Networks for Signal Processing* held on 29th Sept-2nd October 1991 at Nassau Inn, Princeton, NJ, USA.

10. D Hush, Recursive Neural Networks for Signal Processing & Control, *Proc of IEEE Workshop on Neural Networks for Signal Processing* held on 29th Sept-2nd October 1991 at Nassau Inn, Princeton, NJ, USA.

11. P McKee, Signal Identification by Neural Networks, *Proc of IEEE Workshop on Neural Networks for Signal Processing* held on 29th Sept-2nd October 1991 at Nassau Inn, Princeton, NJ, USA.

12. S S Rao, A Neural Network Preprocessor for Multilayer Detection and Estimation, *Proc of IEEE Workshop on Neural Networks for Signal Processing* held on 29th Sept-2nd October 1991 at Nassau Inn, Princeton, NJ, USA.

13. J W Watterson, An Optimum Multilayer Perceptron Neural Receiver for Signal Detection, *IEEE Trans on Neural Network*, vol 1, no 4, December 1990.

14. P D Wasserman, *Neural Computing: Theory and Practice*, Van Nostrand & Reinhold, 1989.

P K DASH, P K NANDA
Department of Electrical Engineering,
Regional Engineering College, Rourkela 769 008, India

A K SWAIN
Electrical Engineering Department,
IGIT, Sarang 759 146, India

AND

H P KHINCHA
Department of Electrical Engineering, Indian Institute of Science,
Bangalore 560 001, India.