The Eurasia Proceedings of Science, Technology,
Engineering & Mathematics (EPSTEM)
ISSN: 2602-3199

The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM), 2021

Volume 14, Pages 1-7

IConTech 2021: International Conference on Technology

An Adaptive Sigmoidal Activation Function for Training Feed Forward Neural Network Equalizer

Zohra ZERDOUMI
M'sila University

Fadila BENMEDDOUR
M'sila University

Latifa ABDOU
Batna II University

Djamel BENATIA
Batna II University

Abstract: Feed for word neural networks (FFNN) have attracted a great attention, in digital communication area. Especially they are investigated as nonlinear equalizers at the receiver, to mitigate channel distortions and additive noise. The major drawback of the FFNN is their extensive training. We present a new approach to enhance their training efficiency by adapting the activation function. Adapting procedure for activation function extensively increases the flexibility and the nonlinear approximation capability of FFNN. Consequently, the learning process presents better performances, offers more flexibility and enhances nonlinear capability of NN structure thus the final state kept away from undesired saturation regions. The effectiveness of the proposed method is demonstrated through different challenging channel models, it performs quite well for nonlinear channels which are severe and hard to equalize. The performance is measured throughout, convergence properties, minimum bit error achieved. The proposed algorithm was found to converge rapidly, and accomplish the minimum steady state value. All simulation shows that the proposed method improves significantly the training efficiency of FFNN based equalizer compared to the standard training one.

Keywords: Non linear equalization, Feed for word neural networks (FFNN), Digital communication channels, Adaptive sigmoidal activation function

Introduction

Achieving high data transmission rate is the main objective in wireless communication systems, though they are confronted to channel impairments, which alter the digital signal and causes inter symbol interference (ISI). Equalization is an approach to mitigate channel ISI and recover the transmitted data (Proakis, 2001; Mehmet et al., 2013). Equalization structure based on linear adaptive filters, limit the performance of the system, nonlinear structures are superior to linear ones; in particular, on non-minimum phase or nonlinear channels (Zerguine et al., 2001; Baloch et al., 2012; Mehmet et al., 2013; Sunita et al., 2015). Many researches have revealed that feed for word neural network (FFNN) equalizers (FFNN) can provide better system performance than conventional ones (Amgothu & Kalaichelvi, 2015; Baloch et al., 2012; Corral et al., 2010; Lyu et al., 2015; Sunita et al., 2015; Zerdoumi et al., 2015).

- Selection and peer-review under responsibility of the Organizing Committee of the Conference

Back propagation (BP) training algorithm is a supervised learning method for the (FFNN) (Haykin, S, 1999), based on steepest descent, its major drawback is the convergence rate wich still very slow. Many researches have been reported to the BP algorithm in order to improve its efficiency and convergence capabilities (Saduf, 2013; Wang et al., 2004). An overview of learning strategies in FFNN as well as numerous improvements in steepest descent through BP algorithm was provided in Schmidhuber (2015).

Among BP learning speed up algorithms, those using the adaptation of the activation function plays a decisive role (Daqi et. al., 2003; Chandra et. al., 2004). It has been recognized that training algorithms that adapt the activation function lead to faster training than those that do not (Chandra et. al.,2004; Daqi et. al., 2003; Xu & Zhang, 2001; Yu et al., 2002).

In our approach we propose an improvement to the FFNN learning algorithm by adapting the nonlinear activation function. Then, the proposed learning algorithm is derived to adjust the free parameter as well as the connection weights and bias in the FFNN structure. The proposed method adapts one free adaptive parameter mutually with weights and bias, therefore computational burden is avoided and the convergence capabilities of the algorithm are significantly improved.

Feed forward neural network-based equalizer

The equalization technique is performed at the receiver to compensate channel impairments such as ISI, noise and nonlinearities, therefore the transmitted is recovered. Conventionally equalizers are structured as adaptive digital filters. More recently nonlinear structures based on neural network are used to enhance system performance, since they can perform complex mapping between input and output spaces and are capable of forming nonlinear decision boundaries. Neural network equalizers perform well on severe environment and can deal with nonlinear ISI due to their strong capabilities. In our work we consider only feed forward neural network-based equalizer as depicted in (Figure 1).

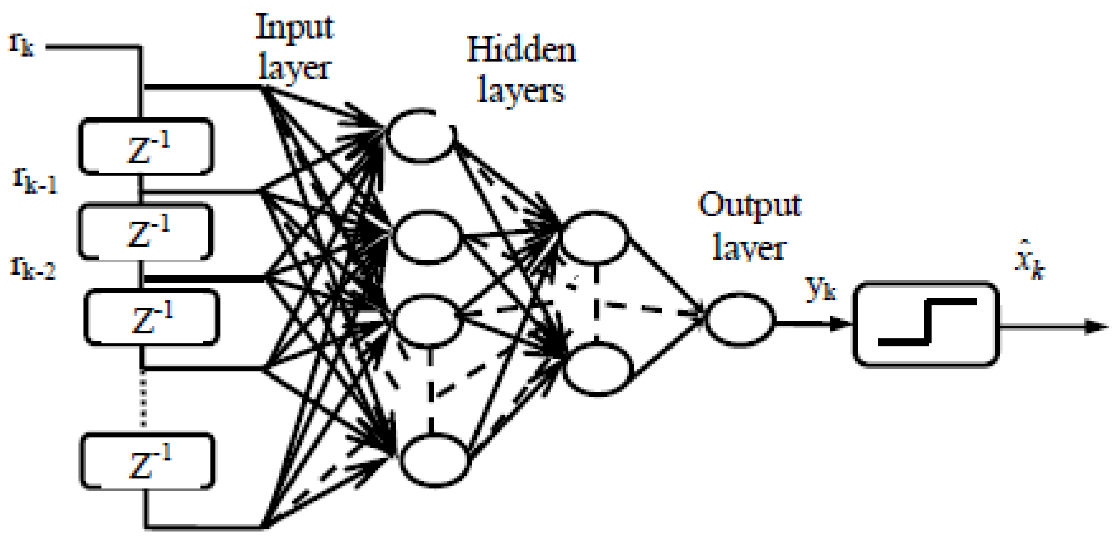

Figure 1. Feed forward neural network-based equalizer

The Feed forward neural network-based equalizer (Figure 1) is realized by connecting the received signal to a time delay line (TDL) which is connect to the input layer of the FFNN. The FFNN is composed by three layers; the input layer connected to the TDL, two hidden layers and an output layer connected to the hard decision device.

Activation functions

Activation functions play an important role in the training procedure of FFNN, they procure nonlinearity and establish the speed convergence of the learning algorithms (Daqi et. al., 2003; Sartaj et. al., 2014). It has been established that adapting the activation function lead to faster training and more flexibility in achieving the nonlinear behavior required (Chandra et. al.,2004; Daqi et. al., 2003). Owing the ability of adapting the activation function the nonlinear capability and the flexibility of the FFNN is improved considerably; consequently, the convergence properties are enhanced. Sigmoidal activation functions are the most commonly used in the FFNN structure. We consider a parametric bipolar-sigmoid (Power, 2001) as the nonlinear activation function. Such function and its derivative are defined as below:

$$\varphi(x) = \frac{1-e^{(-\beta x)}}{1+e^{(-\beta x)}} \; ; \quad \varphi'(x) = 2\beta\left(1+e^{-\beta x}\right)^{-2} e^{-\beta x} = \frac{\beta}{2}\left(1-(\varphi\,(x))^2\right) \tag{1}$$

Where β is the slope parameter, it controls the nonlinearity of the sigmoid. The derivative of $\varphi(x)$ is given by:

Adaptive sigmoidal activation function learning algorithm

FFNN structure is composed by a set of sensorial units organized in hierarchical layers; the input layer, one or more hidden layers and the output layer. The consecutive layers are completely linked. The outputs of the neurons in one layer form the inputs to the next layer. The information is progressed from the input layer to the output layer. The hidden layers perform complex nonlinear mappings between the input and the output layer via the nonlinear activation function (Haykin, S, 1999). In our approach we propose an algorithm that enhances the nonlinear capabilities of the FFNN by adapting the sigmoidal activation function slope of the hidden layers. Thus procures more flexibility and suppleness to the FFNN structure.

Consider a FFNN with an input layer, one hidden layer and one out put layer. Let w_{ik} be the weight that links the inputs to the hidden layer and w_{ji} the weight that links the hidden layer to the output layer. Let $k=\overline{1,m}$, be the inputs index, $i=\overline{1,p}$, the hidden layer index and $j=\overline{1,n}$ the output index. In order to train the sigmoid activation function to perform the desired mapping, a cost function is defined as the sum of the squared error between the actual network output and the desired output, as expressed below:

$$J = \frac{1}{2}\sum_{j=1}^{n}\left(d_j - \mathrm{y}_j\right)^2 \tag{2}$$

We define the following entities for the output's neuron i and hidden layer neuron j respectively as follows:

$$net_i = \sum_{k=1}^{m} \mathrm{w}_{ik} x_k + \theta_i \tag{3}$$

$$net_j = \sum_{i=1}^{p} \mathrm{w}_{ji} y_i + \theta_j \tag{4}$$

The outputs of the neuron i of the hidden layer and the neuron j of the out put layer are given respectively by:

$$\mathrm{y_i} = f(net_i) = \frac{1-e^{-\beta_i net_i}}{1+e^{-\beta_i net_i}} \tag{5}$$

$$\mathrm{y_j} = f(net_j) = \frac{1-e^{-\beta_j net_j}}{1+e^{-\beta_j net_j}} \tag{6}$$

The proposed learning method baptized adaptive activation function (AAF) algorithm adjusts the FFNN common parameters in addition to the sigmoid activation function parameter using the gradient rule. Let's evaluate the gradient of the cost function regarding to the FFNN parameters explicitly weights, bias and sigmoid parameter via considering the output neuron j as flow:

$$\frac{\partial J}{\partial w_{ji}} = \frac{\partial J}{\partial y_j} \cdot \frac{\partial y_j}{\partial net_j} \frac{\partial net_j}{\partial w_{ji}} = \frac{\beta_j}{2}\left(y_j - d_j\right)\left(1 - y_j^2\right)y_i \tag{7}$$

$$\frac{\partial J}{\partial \theta_j} = \frac{\partial J}{\partial y_j} \cdot \frac{\partial y_j}{\partial \theta_j} = \frac{\beta_j}{2}\left(y_j - d_j\right)\left(1 - y_j^2\right) \tag{8}$$

$$\frac{\partial J}{\partial \beta_j} = \frac{\partial J}{\partial y_j} \cdot \frac{\partial y_j}{\partial \beta_j} = \left(y_j - d_j\right)\frac{net_j}{2}\left(1 - y_j^2\right) \tag{9}$$

In the same manner as above, we evaluate the gradient of the cost function regarding to the FFNN parameters by considering the hidden neuron i :

$$\frac{\partial J}{\partial w_{ik}} = \frac{\partial J}{\partial net_i} \cdot x_k = \frac{\beta_i}{2}\left(\sum_{j=1}^{n} \frac{\beta_j}{2}\left(y_j - d_j\right)\left(1 - y_j^2\right)w_{ji}\right)\left(1 - y_i^2\right)x_k \tag{10}$$

$$\frac{\partial J}{\partial \theta_i} = \frac{\partial J}{net_i} \cdot \frac{net_i}{\partial \theta_i} = \frac{\beta_i}{2}\left(\sum_{j=1}^{n} \frac{\beta_j}{2}\left(y_j - d_j\right)\left(1 - y_j^2\right)w_{ji}\right)\left(1 - y_i^2\right) \tag{11}$$

$$\frac{\partial J}{\partial \beta_i} = \frac{\partial J}{\partial y_i} \cdot \frac{\partial y_i}{\partial \beta_i} = \left(\sum_{j=1}^{n} \frac{\beta_j}{2}\left(y_j - d_j\right)\left(1 - y_j^2\right)w_{ji}\right)\frac{net_i}{2}\left(1 - y_i^2\right) \tag{12}$$

Therefore, the parameters on the FFNN are adjusted through the gradient descent rule when the output neuron j is considered as:

$$w_{ji}(n+1) = w_{ji}(n) - \eta\frac{\beta_j}{2}\left(y_j - d_j\right)\left(1 - y_j^2\right)y_i \tag{13}$$

$$\theta_j(n+1) = \theta_j(n) - \eta\frac{\beta_j}{2}\left(y_j - d_j\right)\left(1 - y_j^2\right) \tag{14}$$

$$\beta_j(n+1) = \beta_j(n) - \eta\left(y_j - d_j\right)\frac{net_j}{2}\left(1 - y_j^2\right) \tag{15}$$

When the hidden neuron i is considered, the parameters on the FFNN are adjusted as:

$$w_{ik}(n+1) = w_{ik}(n) - \eta\frac{\beta_i}{2}\left(\sum_{j=1}^{n} \frac{\beta_j}{2}\left(y_j - d_j\right)\left(1 - y_j^2\right)w_{ji}\right)\left(1 - y_i^2\right) \cdot x_{ki} \tag{16}$$

$$\theta_i(n+1) = \theta_i(n) - \eta\frac{\beta_i}{2}\left(\sum_{j=1}^{n} \frac{\beta_j}{2}\left(y_j - d_j\right)\left(1 - y_j^2\right)w_{ji}\right)\left(1 - y_i^2\right) \tag{17}$$

$$\beta_i(n+1) = \beta_i(n) - \eta\left(\sum_{j=1}^{n} \frac{\beta_j}{2}\left(y_j - d_j\right)\left(1 - y_j^2\right)w_{ji}\right)\frac{net_i}{2}\left(1 - y_i^2\right) \tag{18}$$

The adaptive activation-based algorithm is considered as an improvement of the conventional back propagation it includes the traditional fixed activation function as a particular case therefore, it provides more suppleness and nonlinearity to the MLP structure. The proposed algorithm has a higher probability of not getting stuck in local minima. This is principally due to the effect of change in the value of the slope parameter.

Results and Discussion

In our simulations we use an FFNN structure with a single hidden layer and a single node in the output layer. We perform several simulations to reveal the ability and the convergence properties provided by the proposed approach. All simulation results were realized using MATLAB.

Channel Model

We adopt nonlinear channel equalization problems, widely used in literature to evaluate the performance of the equalizers (Corral et al., 2010; Zerguine et al., 200). Nonlinear channel models (NCh_1 and NCh_2) are composed by a linear channel (Ch1 and Ch2) followed by a memory less nonlinearity. The linear channels are represented by their impulses responses coefficients by the following equations:

$$Ch_1 = [0.6963\ \ 0.6964\ \ 0.1741] \tag{19}$$

$$Ch_2 = [0.2600\ \ 0.9300\ \ 0.2600] \tag{20}$$

Ch_1 is a linear minimum phase channel where Ch_2 is a linear non minimum phase channel. The nonlinearity is of polynomial type, described by the above equation:

$$v_k = a_1 u_k + a_2 u_k^2 + a_3 u_k^3 \tag{21}$$

The linear channel output is u_k , whereas v_k is the output of the memory less nonlinearity. Coefficients a_i are scalars, which control the nonlinearity degree (Corral et al., 2010; Zerguine et al., 200). Parameters, a_1, a_2, and a_3 of (21) are set to 1, 0.2, and -0, 1 respectively as given in (Corral et al., 2010).

Performance's Measure

Mean square error convergence

Figure 2 illustrates the convergence behavior of the FFNN for the BP and BPAAF algorithms considering nonlinear channels NCh1and NCh_2. The proposed algorithm BPAAF shows a clear improvement in the convergence time and the steady state value of averaged square error produced than the BP. Despite the nonlinearity the BPAAF achieves the best performance in steady state MSE, it achieves -32dB and -30dB for Nch_1 and Nch_2 respectively. While the steady state MSE reached by the BP is about -25dB and -22dB for Nch_1 and Nch_2 respectively. Thus resulting of a gain of 7dB and 8dB over the BP for Nch_1 and Nch_2 respectively. These improvements are summarized in Table1.

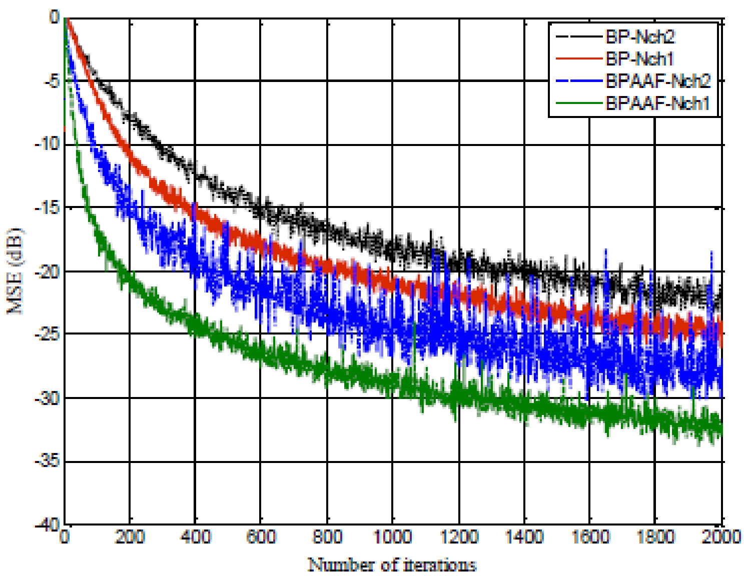

Figure 2. Convergence of the algorithms for nonlinear channels

Bit error rate Study

Figures 3 illustrate the BER curves for the BPAAF and the BP algorithms through nonlinear channels $NCh_{1,}$ $_{NCh2}$. The BPAAF still consistently behaving better than the BP; it shows a gain of about 1.8 dB and 1.7dB over the BP at a BER of 10^{-3}. We can also notice that the BPAAF achieves the minimum BER is about $10^{-3.4}$ at the SNR of 14 dB where the BP realizes a BER of $10^{-2.7}$ at the same SNR for the nonlinear channel $NCh_{1.}$

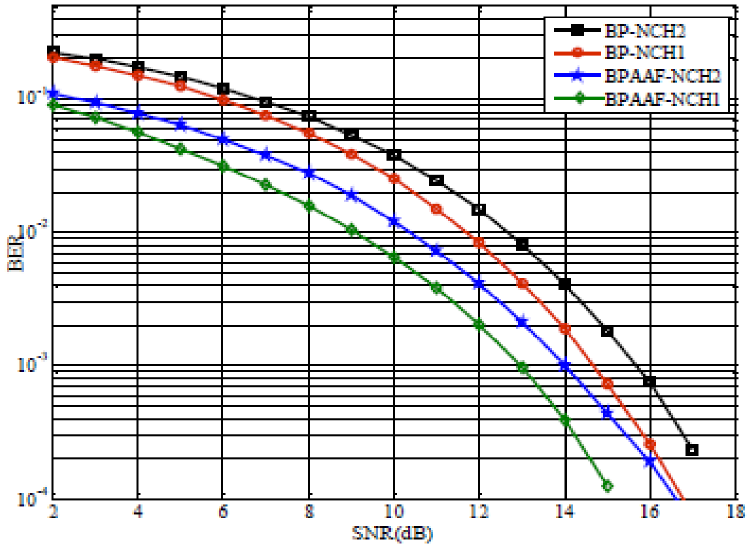

Figure 3. BER curves of the algorithms for nonlinear channels NCh_1, NCh_2

Table 1. Performance analysis of the proposed algorithm (BPAAF)

Channels	Nature	Convergence time of BPAAF over BP (iterations)	Steady state MSE of BPAAF over BP (dB)	SNR at 10^{-3} BER of BPAAF over BP (dB)
Ch1	Linear	440	7.3	1.9
Ch2	Linear	840	8	1.6
Nch1	Nonlinear	602	7	1.8
Nch2	Nonlinear	950	8	1.7

Conclusion

We proposed an algorithm that adapts the sigmoid activation function slope for the FFNN based equaliser. Our approach performs quite well for on nonlinear channels which are difficult to equalize. The proposal BPAA manifests a fast convergence and a lower steady state MSE than the BP. For the BER performance the BPAAF accomplishes all the time the minimum BER. It can be seen from the entire scenarios presented that the BPAAF achieves accurately the best performance. Simulations illustrate that as the severity of the channel increases, the steady state error of BP and BPAA also increase. However, the BPAAF all the times holds lower steady state error.

Scientific Ethics Declaration

The authors declare that the scientific ethical and legal responsibility of this article published in EPSTEM journal belongs to the authors.

References

Chandra, P., & Singh, Y. (2004). An activation function adapting training algorithm for sigmoidal feedforward networks. *Neurocomputing*, *61*, 429-437.

Corral, P., Ludwig, O., & Lima, A. D. C. (2010). Time-varying channel neural equalisation using Gauss-Newton algorithm. *Electronics Letters*, *46*(15), 1055-1056.

Daqi, G., & Genxing, Y. (2003). Influences of variable scales and activation functions on the performances of multilayer feedforward neural networks. *Pattern Recognition*, *36*(4), 869-878.

Haykin, S. (1999). *Neural networks: A comprehensive foundation*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall

Lyu, X., Feng, W., Shi, R., Pei, Y., & Ge, N. (2015, April). Artificial neural network-based nonlinear channel equalization: A soft-output perspective. In *2015 22nd International Conference on Telecommunications (ICT)* (pp. 243-248). IEEE.Proakis, J.G.& Salehi.M. (2008) *Digital communications* (5^{th} Ed). McGraw-Hill.

Saduf, M. A. W. (2013). Comparative study of back propagation learning algorithms for neural networks. *International Journal of Advanced Research in Computer Science and Software Engineering*, *3*(12), 1151-1156.

Schmidh ber, J. (2015). 'Deep learning in neural networks n overview'.*Neural Networks*, *vol.61*, pp. 85–117.

Wang, X., Tang, Z., Tamura, H., Ishii, M., & Sun, W. D. (2004). An improved backpropagation algorithm to avoid the local minima problem. *Neurocomputing*, *56*, 455-460.

Yu, C. C., Tang, Y. C., & Liu, B. D. (2002, October). An adaptive activation function for multilayer feedforward neural networks. In *2002 IEEE Region 10 Conference on Computers, Communications, Control and Power Engineering. TENCOM'02. Proceedings.* (Vol. 1, pp. 645-650). IEEE.

Zerdoumi, Z., Chicouche, D., & Benatia, D. (2015). Neural networks based equalizer for signal restoration in digital communication channels. *International Letters of Chemistry, Physics and Astronomy*, *55*(1), 191-204.

Zerdoumi, Z., Chikouche, D., & Benatia, D. (2016). Multilayer perceptron based equalizer with an improved back propagation algorithm for nonlinear channels. *International Journal of Mobile Computing and Multimedia Communications (IJMCMC)*, *7*(3), 16-31.

Zerdoumi, Z., Chikouche, D., & Benatia, D. (2016). An improved back propagation algorithm for training neural network-based equaliser for signal restoration in digital communication channels. *International Journal of Mobile Network Design and Innovation*, *6*(4), 236-244.

Zerguine, A., Shafi, A., & Bettayeb, M. (2001). Multilayer perceptron-based DFE with lattice structure. *IEEE transactions on neural networks*, *12*(3), 532-545.

Author Information

Zohra ZERDOUMI
Electronic department, laboratory of electrical engineering LGE, M'sila University
University of Med BOUDIAF-BP 166 M'sila 28000, Algeria
Contact e-mail : zohra.zerdoumi@univ-msila.dz

Fadila BENMEDDOUR
Electronic department, laboratory of electrical engineering LGE, M'sila University
University of Med BOUDIAF-BP 166 M'sila 28000, Algeria

Latifa ABDOU
Electronic Department**,** Batna II
Avenue Chahid Boukhlouf 05000, Algeria

Djamel BENATIA
Electronic Department, Batna II University,
Avenue Chahid Boukhlouf 05000, Algeria

To cite this article:

Zerdoumi, Z., Benmeddour, F., Abdou, L., & Benatia, D. (2021). An adaptive sigmoidal activation function for training feed forward neural network equalizer. *The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM), 14*, 1-7.