COMPLEX NEURAL NETWORKS FOR EQUALIZING NONLINEAR DIGITAL RADIO LINKS

N. Benvenuto*, M. Marchesi**, F. Piazza, A. Uncini

Dipartimento di Elettronica e Automatica, Università di Ancona
Via Brecce Bianche, 60131 Ancona - Italy

* Dipartimento di Elettronica ed Informatica, Università di Padova
Via Gradenigo, 6A, 35131 Padova - Italy

** D.I.B.E., Università di Genova
Via all’Opera Pia, 11A, 16145 Genova - Italy

ABSTRACT

In the last few years, due to their capabilities of classification, generalization, and learning-from-examples, neural network models have been used in many DSP applications. However, in some DSP problems the available signals and desired outputs are in complex form; in these cases, many widely-used neural models cannot be directly applied because they require both input and target signals to be real-valued.

In this paper, a novel Multi Layer Perceptron (MLP) structure is proposed which is able to process complex signals. Since the classical real-valued sigmoidal activation function, analytically extended to the complex plane, is not bounded everywhere (due to the Liouville’s theorem), a new complex activation function is proposed which is always bounded and behaves in a way very close to its real-valued counterpart. However, this function is not analytic and the standard Back-Propagation (BP) learning algorithm cannot be used. Therefore, for complex MLP networks employing neurons with such activation function, a new BP learning algorithm is presented. This is very similar to the real-valued BP and collapses to it when all the involved quantities become real-valued.

The proposed network has been applied to the problem of equalizing the nonlinearities that appear in digital radio links carrying M-level quadrature amplitude modulation (M-QAM), when the high-power amplifier (HPA) is operated near its maximum output power (as in satellite systems). The obtained results, compared with the complex Adaptive Linear Combiner (ALC) applied to the same problem, demonstrate the capabilities of the proposed approach.
1. INTRODUCTION

Neural Network models are increasingly used in DSP applications, owing to their capabilities of generalization and learning from examples, their robustness, and their speed (due to the intrinsic parallelism of these models [1]). In source coding (vectorial quantization), Kohonen unsupervised nets are preferred [2],[3]. In non-linear digital filtering and equalization instead the most used neural model is the Multi-Layer Perceptron (MLP) [4]. MLP is a very general structure, capable to process data with high parallelism, and is provided of a powerful gradient-based learning algorithm, called Back-Propagation (BP).

Data processed by MLPs are in general real-valued, and the BP algorithm has been developed for such data. However, in some DSP problems the input and output signals are in complex form. For instance, in some communication systems a complex envelope representation is used instead of the more cumbersome real-valued bandpass. In such cases, it is possible to use a classical real-valued MLP where the complex input and output signals of the network are replaced by pairs of independent real-valued signals, representing the real and imaginary part of the complex signal, or its modulus and its phase. However, this approach does not make use of the intrinsic structure of the complex signals.

Another approach, which will be followed in this paper, is the development of a complex version of BP algorithm able to deal with complex-valued neurons, inputs and outputs. To be successful, such an approach must retain the power and the simplicity of the recursive notation of the real BP, and it must reduce to the classical real-valued BP if the imaginary part of the involved quantities is set to zero. Moreover, on problems concerning complex-valued inputs and outputs, the complex BP must perform better than its real-valued counterpart.

In the literature, there are some papers dealing with complex BP [5,6]. Our approach, however, is different because we preferred to avoid the use of analytic neuron activation functions. In fact, due to the Liouville's theorem, non-constant analytic complex functions are unbounded for some points in the complex plane. This property contrasts with the behaviour of real activation functions, such as sigmoid and tan\(^{-1}\) functions, which usually are monothone, bounded and very smooth. The preservation of such smoothness, which greatly enhances BP convergence properties, is in our opinion more important that having an analytic activation function. The drawback of this approach is that in this way it is not possible to directly apply the equations of the classical BP. We shall show, however, that it is possible to overcome this problem by developing learning equations that are similar (but not equal) to the real-valued BP equations and retain all their properties. In particular they collapse to real-valued BP if applied to real-valued signals and network parameters.
In this paper, the new complex MLP is applied to the equalization of digital radio links carrying M-level quadrature amplitude modulation (M-QAM). In particular those found in satellite transmission systems, when the high-power amplifier (HPA) is operated near its maximum output power. In this situation, nonlinearities that can cause serious performance degradation are introduced and a channel equalization becomes necessary. Here, the number of modulation levels is 64, and the channel has been simulated following the model of [7], with some simplifications. The proposed complex-valued MLP algorithm has been applied to the problem of equalizing the HPA nonlinearities at the receiver, using a training sequence. The obtained results, compared with the complex Adaptive Linear Combiner (ALC) applied to the same problem, demonstrate the capability of the proposed approach. This result confirms an analogous result, found on another equalization problem dealing with PSK/TDMA satellite channels [8].

2. COMPLEX BACK-PROPAGATION

The proposed MLP has neurons whose inputs, outputs and weights are complex-valued. Each neuron computes the (complex) weighted sum of its inputs, and calculates the new activation value.

Forward Phase: for \( s = 1, ..., M \) and \( k = 1, ..., N_s \),

\[
net_k(s) = \sum_{j=0}^{N_{s-1}} w_{kj}^{(s)} o_j^{(s-1)},
\]

\[ o_k^{(s)} = F(\text{net}_k^{(s)}) , \tag{1} \]

where "s" is the layer index, \( s=0 \) is the input layer, \( s=1,...,M-1 \), are the hidden layers, and \( s=M \) is the output layer], "k" is the neuron index [\( k=0 \) refers to bias neurons with output set to 1, while \( k=1,...,N_s \), refer to neurons of layer \( s \)], \( o_k^{(s)} \) is the activation output of \( k \)-th neuron of \( s \)-th layer, and \( w_{kj}^{(s)} \) is the synaptic weight of \( k \)-th neuron of \( s \)-th layer, relative to the \( j \)-th neuron of \( (s-1) \)-th layer. Notice that the weight \( w_{k0}^{(s)} \) acts as an offset for the \( k \)-th neuron of \( s \)-th layer.

Let \( z = x + jy \), where "j" is the imaginary unit while \( x \) and \( y \) are real-valued variables, the neuron complex activation function, \( F(z) \), is derived from the classical sigmoidal real function \( f(x) \)

\[
f(x) = \frac{2}{1 + e^{-C_2 x}} - C_1 , \tag{3} \]

as follows:
F(z) = F(x + jy) = f(x) + j f(y)

with $C_1, C_2$ suitable constants. It is seen that $F(z)$ is bounded everywhere in the complex plane and collapses to $f(x)$ if it is restricted to real axis. However, $F(z)$ is not analytic.

With such assumptions on the complex neuron activation function, it is possible to derive an iterative learning algorithm very similar to the real-valued classical BP. By defining a quadratic error function, $E$, as the sum of the squared modules of the differences between desired and computed MLP outputs (as in the classical BP), the application of the steepest descend method brings to the following equations:

**Learning Phase:** for $s = M, ..., 1$ and $k = 1, ..., N_s$

$$
\sigma_k^{(s)} = \begin{cases} 
  t_k - \sigma_k^{(M)}, & \text{for } s = M \\
  \sum_{j=1}^{N_{s+1}} w_{jk}^{*(s+1)} \delta_j^{(s+1)}, & \text{for } s = M-1, ..., 1
\end{cases}
$$

with

$$
\delta_k^{(s)} = \text{Re}[\sigma_k^{(s)}] f'(\text{Re}[\text{net}_k^{(s)}]) + j \text{Im}[\sigma_k^{(s)}] f'(\text{Im}[\text{net}_k^{(s)}]),
$$

and the weight adjustments are:

$$
\Delta w_{kj}^{(s)} = -\eta \delta_k^{(s)} o_j^{*(s-1)}, \quad \text{for } j = 0, ..., N_{s-1}
$$

where (*) denotes complex conjugate, $t_k, k = 1, ..., N_M$, is the target output, $f'(.)$ is the derivative of $f(.)$, $\eta$ is a suitable positive real constant (learning rate), and $\Delta w_{kj}^{(s)}$ is the computed synaptic weight variation. It is worth noting that equations (5)-(7) are not equal to the real-valued BP equations, but collapse to them if all involved complex quantities become real-valued.

### 3. MODEL OF THE CHANNEL

The digital radio link shown in Fig. 1, is a simplified version of the model reported in [7]. In practice, the nonlinearities are due to the HPA, while the radio channel is considered ideal and without noise.
The input data stream is a sequence of M-level complex values \( \{a_n\} \). Admissible values for such data are complex numbers having real and imaginary part belonging to the ensemble: \{-7, -5, -3, -1, 1, 3, 5, 7\}, with cardinality equal to 64. Input data are spaced by \( T \) seconds.

The input stream \( \{a_n\} \) is then fed into a pulse shaping circuit characterized by an impulse response \( g(t) \). Its transfer function is bandlimited and has a Fourier transform \( G(f) \) given by the formula:

\[
G(f) = T \sqrt{C(f, \alpha)}
\]

where \( C(f, \alpha) \) is the cosine-rolloff function typically used in digital radio [7]. Here the rolloff factor \( \alpha \) is equal to 0.5. After pulse shaping, the complex input sequence is transformed into the continuous time signal \( u(t) \):

\[
u(t) = \sum_n a_n g(t - nT)
\]

(9)

Since the analog predistorter is not present in our case, the signal \( u(t) \) is fed directly into the HPA which is operating near its maximum power. \( u(t) \) is distorted according to the formula obtained from [7, equation (3)] for \( P_{\text{max}} = 1 \) and \( \Phi_0 = \frac{\pi}{6} \). The result is the signal:

\[
w(t) = \frac{2}{1 + |u(t)|^2} u(t) e^{j \frac{\pi}{6} \frac{2 |u(t)|^2}{1 + |u(t)|^2}}
\]

(10)

This signal is then transmitted through an ideal noiseless radio channel. The transfer function of the receiver is the same of the pulse shaping circuit, \( G(f) \). After this block, the output signal, \( x(t) \), is fed into the complex MLP equalizer. To be processed by the neural network, the received continuous signal \( x(t) \) has to be sampled. In order to maintain the spectral characteristics of \( x(t) \), this is accomplished by sampling \( x(t) \) with a period \( T' \), \( \kappa \) times smaller than \( T \). The discretized signal, \( \{x_i\} \), is therefore given by:

\[
x_i = x(i T'), \quad T' = \frac{T}{\kappa}
\]

(11)

The proposed neural equalizer has some complex inputs, one hidden layer and an output layer. This layer is composed of only a single linear neuron, having a linear activation function \( f(x) = Kx \). In order to correctly reproduce the signal level range at the output, this neuron is a linear combiner. The output sequence of the digital link, \( \{x_i\} \), is presented to the inputs of the MLP, and is shifted by \( \kappa \) samples at each iteration. The performance criterion is the mean-square-error between the network output at time \( n \), \( \alpha_{1,n}(M) \), and the target output which is equal to the corresponding input M-level sample, \( a_n \).
4. SOME EXPERIMENTAL RESULTS

In order to evaluate the performance of the proposed equalization method, the digital radio link of Fig.1 was simulated on a computer, for an oversampling parameter $\kappa$ of 5 and a maximum instantaneous power of $u(t)$ of -2 dB. Using an input data sequence $\{a_n\}$ of R=194 symbols (corresponding to 970 signal samples) as learning set, three different experiments have been carried out.

In the first experiment, a complex MLP with 21 complex inputs and 5 hidden neurons was trained with the learning set until the minimum mean-square-error (MSE) at the output was obtained. The MSE was defined to be

$$\text{MSE} = \frac{1}{R} \sum_{n=0}^{R-1} |o_{1,n}^{(M)} - a_n|^2$$

(12)

The $C_1$ and $C_2$ parameters of the sigmoid function $f(x)$ were chosen to be 2.164 and 1.0 respectively, while the learning rate $\eta$ was adaptively adjusted. Fig. 3a reports the sequence at the baud rate, projected onto the first quadrant of the complex plane, obtained before the equalizer; the bold black circles indicate the undistorted data. On the other hand, Fig.3b reports the sequence obtained from the neural equalizer. Notice that now the samples are much more clustered near the correct data than before.

The second experiment was performed for an higher number of hidden neurons (equal to 9). The result obtained in this case is reported in Fig. 3c which shows how this structure can exhibit better performance with respect to the previous one.

In order to compare the proposed approach with the classical ALC, the last experiment was performed using as equalizer a complex adaptive linear combiner with 21 inputs. The result is reported in Fig. 3d, showing that the ALC performs much worse than the complex MLP. The results obtained in all cases are summarized, in term of MSE, in Table 1, where are also reported the number of presentation of the whole learning set to the networks (epochs).
Fig. 1 Block diagram of the digital radio link and the complex MLP equalizer.
REFERENCES


---

TABLE 1. Number of free parameters, performed epochs and obtained MSE's for the complex linear combiner with 21 inputs (C21_1), the complex MLP with 21 inputs and 5 hidden neurons (C21_5_1), and the complex MLP with 21 inputs and 9 hidden neurons (C21_9_1).

<table>
<thead>
<tr>
<th>Net.</th>
<th>n. complex param.</th>
<th>n. of epochs</th>
<th>MSE [dB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>C21_1</td>
<td>22</td>
<td>20000</td>
<td>-22.69</td>
</tr>
<tr>
<td>C21_5_1</td>
<td>116</td>
<td>20000</td>
<td>-24.50</td>
</tr>
<tr>
<td>C21_9_1</td>
<td>208</td>
<td>20000</td>
<td>-27.08</td>
</tr>
</tbody>
</table>
Fig. 2 Sequence obtained:
(a) at the input of the equalizer;
(b) at the output of the equalizer when the complex MLP has 21 inputs and 5 hidden neurons;
(c) when the complex MLP has 21 inputs and 9 hidden neurons;
(d) when the complex linear combiner has 21 inputs.
The correct symbol positions are indicate by the bold black circles.