IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received July 2, 2019, accepted July 16, 2019, date of publication July 17, 2019, date of current version August 5, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2929590

Butterfly Neural Filter Applied to Beamforming

TIAGO F. B. DE SOUSA[1] AND MARCELO A. C. FERNANDES[2]
[1]Federal Institute of Rio Grande do Norte, Ceará-Mirim 59580-000, Brazil
[2]Department of Computer Engineering and Automation, Federal University of Rio Grande do Norte (UFRN), Natal 59064-741, Brazil

Corresponding author: Marcelo A. C. Fernandes (mfernandes@dca.urn.br)

This work was supported in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)—Finance Code 001.

ABSTRACT The butterfly neural beamformer (NB-Butterfly) is a new adaptive multiple-antenna spatial neural filter inspired on the neural butterfly equalizer (NE-Butterfly), a filter intended to equalize any channel that has real or complex taps, whether linear or nonlinear. Due to the broad use cases of the NE-Butterfly, the objective in this paper is to introduce this novel beamforming filter, the NB-Butterfly and analyze its performance by comparing to other neural and linear beamformers, while also presenting an enhanced training strategy that wasn't present in the butterfly neural architecture before, which is called butterfly neural beamformer with joint error (NB-Butterfly-JE). The proposals are evaluated and compared for different types of channels in order to validate their performance in different use cases.

INDEX TERMS Butterfly beamformer, adaptive beamforming, artificial neural networks, neural beamformer.

I. INTRODUCTION

Spatial filters with adaptive multiple-antenna processing are known to dramatically enhance wireless communication systems, being able to identify signals transmitted on the same carrier frequency that are sufficiently separated in the spatial domain [1], [2]. In these systems, one of the main problems that cause signal degradation, besides the white Gaussian noise, is the Inter Symbol Interference (ISI), that happens due to band limitation as well as multi-path effects, which causes superposition of symbols of the same transmitter in a receiver [3].

The technique most commonly used in the processing of adaptive antennas is known as beamforming, where the general algorithm creates a linear combination of the received signals at different elements in an antenna array. Traditionally, these algorithms use the minimum mean square error (MMSE) principle to achieve the desired output through the use of a known reference signal and the output of the filter, however when considering space-division multiple access (SDMA), where channel configuration and spatial separation in angles of arrival (AOA) of the signal and other sources dictates the performance of the communication, what happens when the conditions aren't favorable is that the individual sources become linearly inseparable [4]–[7]. In order to be able to reconstruct the desired signal, a nonlinear beamforming system is necessary, which can be achieved with the use of artificial neural networks (ANN), among other strategies [8]–[11].

The use of ANN-based solutions to beamforming has been quite explored in literature, where [12]–[14] present implementations and strategies on where each ANN architecture is best used in order to solve the problem. More studies on this area present different ways to solve this problem, such as in [15], where a comparison between different complex-valued radial basis function (RBF) ANNs are made. In [16] a RBF ANN was used in order to implement beamforming in one or two dimensional antenna arrays, which had very good results in comparison to the Wiener solution and [17] applies neural adaptive beamforming (NAB) to speech recognition, where it is used to re-estimate spatial filter coefficients, managing to effectively lower the computational cost compared to what was previously in place.

In this work, a new NAB strategy, called Butterfly Neural Beamformer (NB-Butterfly), is implemented, inspired by the neural equalizer structure in [18], the Butterfly Neural Equalizer (NE-Butterfly), which is composed by four multilayer perceptron (MLP) neural networks trained with the backpropagation algorithm and has the purpose of equalizing channels of any kind, be it with real or complex taps, be it linear or nonlinear, where it performed better than other neural equalizers for different channel conditions. There is also a new training strategy created for the NB-Butterfly, which is based on the joint error strategy presented in [19], being

The associate editor coordinating the review of this manuscript and approving it for publication was Ivan Wang-Hei Ho.

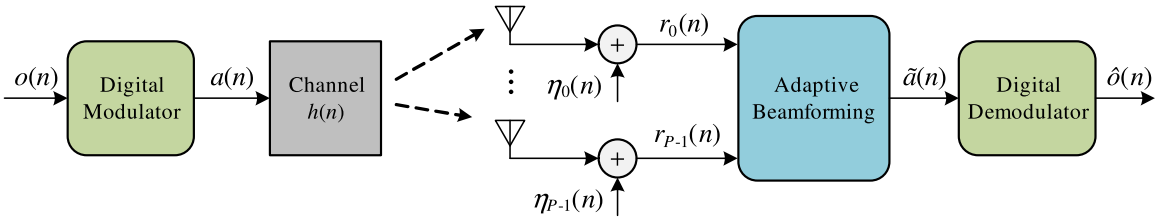


FIGURE 1. Baseband discrete communication system modeled.

named the NB-Butterfly with Joint Error (NB-Butterfly-JE), that aims to speed up the convergence of the backpropagation algorithm present in each individual network. The objective then is to analyze the performance of the NB-Butterfly and the NB-Butterfly-JE when compared to other NABs and linear beamformers and evaluate if it validates a new case of use for the MLP Butterfly Neural Network architecture.

The results are presented using BER and mean square error (MSE) curves for the backpropagation training. In order to validate the NB-Butterfly and NB-Butterfly-JE, they were analyzed and compared with another neural beamforming strategy called the Bi-dimensional Neural Beamformer with Joint Error (BNB-JE), which is an implementation based on the Multi Layer Perceptron Bi-dimensional Neural Equalizer with Backpropagation (BNE-MLP-BP) presented in [18]–[22] and a least mean square (LMS) adaptive beamformer presented in [23]. The digital communication system where the beamformers were tested uses 4-QAM (quadrature amplitude modulation), which was simulated for four different channels that have varying levels of multi pathing and AOA between the source signal and its reflections, leading to linear and nonlinear configurations.

II. COMMUNICATION SYSTEM MODEL

Since one of the main problems that affect communication systems, ISI, happens due to the superposition of symbols of the same transmitter in a receiver, which is caused due to band limitation and multi-path effects, The multi-path phenomena are caused due to signal reflexions that occur during its transmission, creating a spreading of the signal in space and time, which is what originates the ISI [1], [3].

Figure 1 shows the baseband discrete communication system modeled where the n-th binary word, $o(n)$, composed of b bits, is mapped in a n-th complex symbol $a(n) = a^I(n) + ja^Q(n)$ by a M-ary digital modulator scheme such as BPSK, QPSK, M-QAM or similars. M represents the number of symbols associated with a digital modulation scheme. The symbols are transmitted through the multi-path channel expressed as

$$h(n) = \sum_{k=0}^{L-1} \alpha_k \delta(n - \tau_k) \tag{1}$$

where L is the number of the paths of the channel, $h(n)$, α_k is complex gain of the k-th path and τ_k is an integer value associated with the delay of the k-th path.

The receiver has the linearly equally-spaced (LES) array composed with P antenna elements arranged along the y-axis, with spacing of Δy. Each p-th antenna element receives the complex signal $r_p(n)$ expressed as

$$\begin{aligned} r_p(n) &= r_p^I(n) + jr_p^Q(n) \\ &= \sum_{k=0}^{L-1} \alpha_k a(n - \tau_k) e^{-j\frac{2\pi p}{\lambda}\Delta y cos(\theta_k)} + \eta_p \end{aligned} \tag{2}$$

where λ is the wavelength, θ_p is p-th AOA of the p-th channel path and η_p is the additive white noise associated with each p-th antenna element. The received signals in all P antenna elements are processed by an adaptive spatial filter with beamforming, also called adaptive beamforming.

The use of adaptive beamforming aims to reduce multi-path effect in current communication systems and the main objective is to control the directionality of the reception or transmission of a signal on an antenna array, in a way so that the it adds the phases of the signals in the desired direction and nulls the unwanted ones. For channels where ISI can be very dynamic, it is necessary to make a beamforming strategy using adaptive algorithms at the receiver. These algorithms conveniently manipulate the gains of the spatial filter, aiming to keep track of a transmitter, optimizing the quality of the received signal [1].

The conventional adaptive beamforming is presented in Figure 2 and the output $\tilde{a}(n)$ can be expressed as

$$\tilde{a}(n) = \sum_{p=0}^{P-1} f_p(n) r_p(n) \tag{3}$$

where the $f_p(n) = f_p^I(n) + jf_p^Q(n)$ is p-th complex gain associated with the p-th antenna element. In each n-th sample the adaptive beamforming updates the P gains using the adaptive algorithm and the error signal $e(n)$ expressed by

$$e(n) = a_{ref}(n) - \tilde{a}(n) \tag{4}$$

where $a_{ref}(n)$ is called the reference signal [1].

As presented in [1], [2], [11], the conventional beamforming is a linear spatial filter, (See Equation 3) and for a LES array with P antennas, the beamforming can insert nulls in $P - 1$ AOAs, in other words, the beamforming scheme can eliminate the ISI for a channel with $P - 1$ interference paths. For the cases where $L > P$ the linear beamforming (LB) tries to minimize the ISI, however the LB performance is deeply dependent of the channel parameters (α, τ and θ).

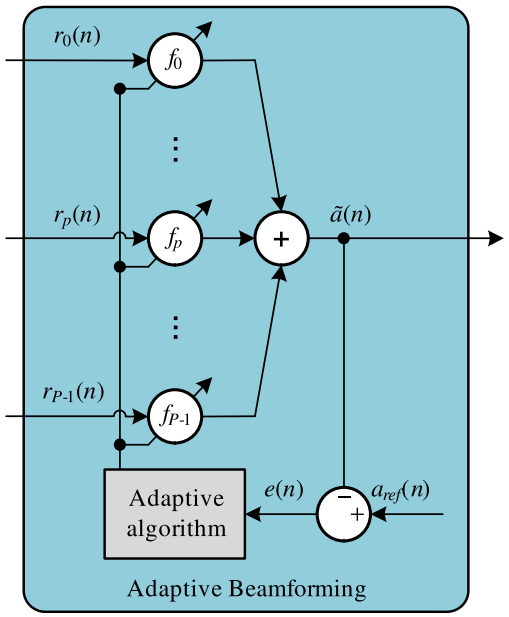


FIGURE 2. Adaptive beamforming scheme.

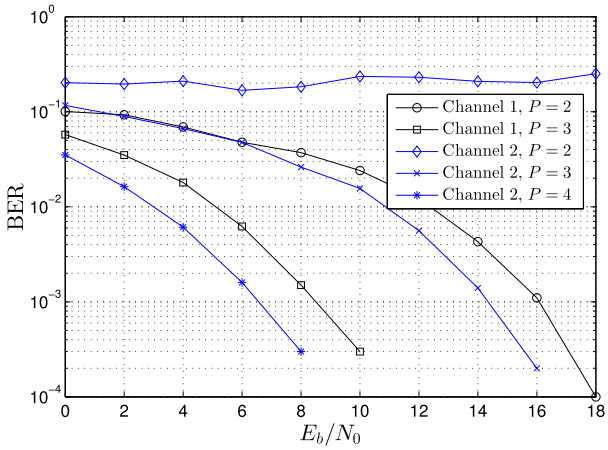


FIGURE 3. BER curve for the channels presented in Table 1. Results were obtained for a communication system (see Figure 1) using BPSK modulation with $P = 2$, $P = 3$, and $P = 4$ antenna elements.

Figure 3 shows the BER curve to the channels presented in Table 1. The curves were obtained for a communication system (see Figure 1) using BPSK ($M = 2$) modulation with $P = 2$, $P = 3$, and $P = 4$ antennas elements. The results show that for channel 1 (see Figure 3), the LB can reduce the BER for all values of P and for channel 2 (see Figure 3) the LB cannot reduce the BER for $P = 2$. The radiation pattern for channel 1 and channel 2 are presented in Figures 4 and 5, respectively.

TABLE 1. Simulated channels used to show the linear classification problem associated with LB.

Channel 1	Path gains	$\alpha_0 = 1.0$	$\alpha_1 = 0.7$	$\alpha_2 = 0.5$	$\alpha_3 = 0.3$
	Path delay	$\tau_0 = 0$	$\tau_1 = 3$	$\tau_2 = 6$	$\tau_3 = 9$
	AOA	$\theta_0 = 0°$	$\theta_1 = 30°$	$\theta_2 = 90°$	$\theta_3 = 120°$
Channel 2	Path gains	$\alpha_0 = 1.0$	$\alpha_1 = 0.9$	$\alpha_2 = 0.8$	$\alpha_3 = 0.5$
	Path delay	$\tau_0 = 0$	$\tau_1 = 3$	$\tau_2 = 6$	$\tau_3 = 9$
	AOA	$\theta_0 = 0°$	$\theta_1 = 30°$	$\theta_2 = 90°$	$\theta_3 = 120°$

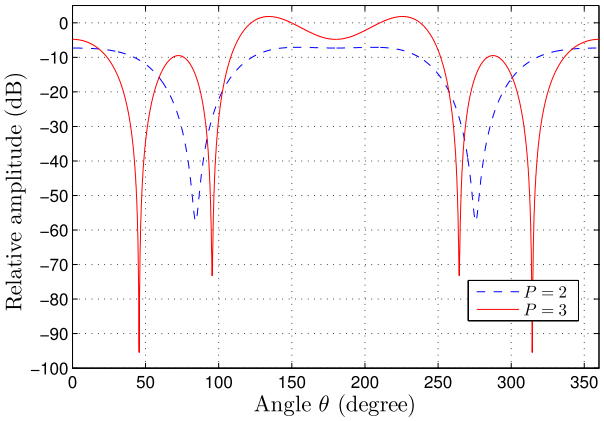


FIGURE 4. LB radiation pattern for channel 1. BPSK system with $P = 2$, and $P = 3$ antennas elements.

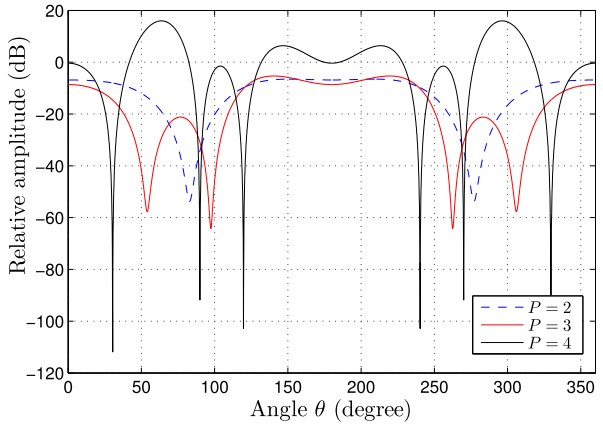


FIGURE 5. LB radiation pattern for channel 2. BPSK system with $P = 2$, $P = 3$, and $P = 3$ antennas elements.

The low performance of the LB to the channel 2 for $P = 2$, it can be explained with a linear classification problem, that is, Equation 2 can be rewrite in terms of channels 1 and 2 ($L = 4$ and $\Delta y = \lambda/2$) as

$$\begin{aligned} r_p(n) &= \sum_{k=0}^{3} \alpha_k a(n-\tau_k) e^{-j2\pi p cos(\theta_k)} + \eta_p \\ &= \alpha_0 a(n-\tau_0) e^{-j\pi p cos(\theta_0)} \\ &\quad + \alpha_1 a(n-\tau_1) e^{-j\pi p cos(\theta_1)} \\ &\quad + \alpha_2 a(n-\tau_2) e^{-j\pi p cos(\theta_2)} \\ &\quad + \alpha_3 a(n-\tau_3) e^{-j\pi p cos(\theta_3)} + \eta_p \end{aligned} \tag{5}$$

using Table 1 values and making $\eta_p = 0$, $r_p(n)$ can be expressed as

$$\begin{aligned} r_p(n) &= \alpha_0 a(n) e^{-j\pi p} + \alpha_1 a(n-3) e^{-j\pi p \times 0.866} \\ &\quad + \alpha_2 a(n-6) e^{-j\pi p \times 0} \\ &\quad + \alpha_3 a(n-9) e^{-j\pi p \times -0.5}. \end{aligned} \tag{6}$$

For $P = 2$ the LB output can be expressed as

$$\tilde{a}(n) = f_0(n) r_0(n) + f_1(n) r_1(n) \tag{7}$$

TABLE 2. All possible values for $r_0^I(n)$, $r_1^I(n)$ and $r_1^Q(n)$ associated with channel 1 and channel 2.

Classes	$a(n)$	$a(n-3)$	$a(n-6)$	$a(n-9)$	channels	$r_0^I(n)$	$r_1^I(n)$	$r_1^Q(n)$
$\tilde{a}(n) < 0$	-1	-1	-1	-1	channel 1	-2.5	1.14	0.01
					channel 2	-3.2	$+1.02$	$+0.13$
	-1	-1	-1	$+1$	channel 1	-1.9	$+1.14$	-0.59
					channel 2	-2.2	$+1.02$	-0.87
	-1	-1	$+1$	-1	channel 1	-1.5	$+2.14$	$+0.01$
					channel 2	-1.6	$+2.62$	$+0.13$
	-1	-1	$+1$	$+1$	channel 1	-0.9	$+2.14$	-0.50
					channel 2	-0.6	$+2.62$	-0.87
	-1	$+1$	-1	-1	channel 1	-1.1	-0.14	$+0.59$
					channel 2	-1.4	-0.62	$+0.87$
	-1	$+1$	-1	$+1$	channel 1	-0.5	-0.14	-0.01
					channel 2	-0.4	-0.62	-0.13
	-1	$+1$	$+1$	-1	channel 1	-0.1	$+0.86$	$+0.59$
					channel 2	$+0.2$	$+0.98$	$+0.87$
	-1	$+1$	$+1$	$+1$	channel 1	$+0.5$	$+0.86$	-0.01
					channel 2	$+1.2$	$+0.98$	-0.13
$\tilde{a}(n) > 0$	$+1$	-1	-1	-1	channel 1	-0.5	-0.86	$+0.01$
					channel 2	-1.2	-0.98	$+0.13$
	$+1$	-1	-1	$+1$	channel 1	$+0.1$	-0.86	-0.59
					channel 2	-0.2	-0.98	-0.87
	$+1$	-1	$+1$	-1	channel 1	$+0.5$	$+0.14$	$+0.01$
					channel 2	$+0.4$	$+0.62$	$+0.13$
	$+1$	-1	$+1$	$+1$	channel 1	$+1.1$	$+0.14$	-0.59
					channel 2	$+1.4$	$+0.62$	-0.87
	$+1$	$+1$	-1	-1	channel 1	$+0.9$	-2.14	$+0.59$
					channel 2	$+0.6$	-2.62	$+0.87$
	$+1$	$+1$	-1	$+1$	channel 1	$+1.5$	-2.14	-0.01
					channel 2	$+1.6$	-2.62	-0.13
	$+1$	$+1$	$+1$	-1	channel 1	$+1.9$	-1.14	$+0.59$
					channel 2	$+2.2$	-1.02	$+0.87$
	$+1$	$+1$	$+1$	$+1$	channel 1	$+2.5$	-1.14	-0.01
					channel 2	$+3.2$	-1.02	-0.13

where

$$r_0(n) = r_0^I(n) = f_0(n)\left(\alpha_0 a(n) + \alpha_1 a(n-3) + \alpha_2 a(n-6) + \alpha_3 a(n-9)\right) \quad (8)$$

and

$$\begin{aligned} r_1(n) &= r_1^I(n) + j r_1^Q(n) \\ &= f_1(n)\Big(\alpha_0 a(n) e^{-j\pi} + \alpha_1 a(n-3) e^{-j\pi \times 0.866} \\ &\quad + \alpha_2 a(n-6) + \alpha_3 a(n-9) e^{-j\pi \times -0.5}\Big). \end{aligned} \quad (9)$$

As $\tilde{a}(n)$ is a BPSK signal, the Equation 7 can be rewrite as

$$\begin{aligned} \tilde{a}(n) &= \Re\{f_0(n) r_0(n)\} + \Re\{f_1(n) r_1(n)\} \\ &= f_0^I(n) r_0^I(n) + f_1^I(n) r_1^I(n) - f_1^Q(n) r_1^Q(n) \end{aligned} \quad (10)$$

where $\Re\{\cdot\}$ returns the real part of the complex number. For BPSK system, there are two classes: $\tilde{a}(n) < 0$ and $\tilde{a}(n) > 0$ and based on Equation 10, it is possible to observer that the hyperplane that classifier the classes ($\tilde{a}(n) = 0$) can be expressed as

$$r_1^Q(n) = \frac{f_0^I(n)}{f_1^Q(n)} r_0^I(n) + \frac{f_1^I(n)}{f_1^Q(n)} r_1^I(n). \quad (11)$$

Table 2 shows all possible values for $r_0^I(n)$, $r_1^I(n)$ and $r_1^Q(n)$, and Figures 6 and 7 plot the values and the hyperplane found by LB (see Equation 11). For channel 1, the LB with $P = 2$ can create a linear classification between two classes,

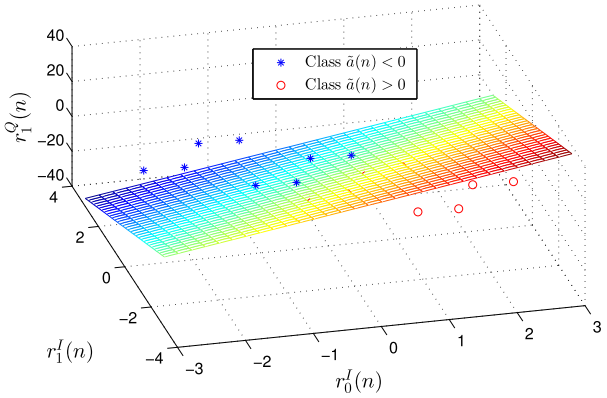


FIGURE 6. Hyperplane $r_1^Q(n) = 5.2 r_0^I(n) - 5.1 r_1^I(n)$ found by LB for channel 1.

Figure 6; however, for channel 2, it is not possible to create a linear classification with $P = 2$, Figure 7 and this deems the utilization of the LB unfeasible. In order to minimize this problem, this work proposes a non-linear classifier called Butterfly Neural Beamformer (NB-Butterfly).

III. BUTTERFLY NEURAL BEAMFORMER

The use of butterfly equalization structures in optical equalizers were the inspiration for the creation of the NE-Butterfly, with this architecture bringing advantages to equalizers, the next step was to investigate whether this is also a viable strategy to filter channel degradation and distortion

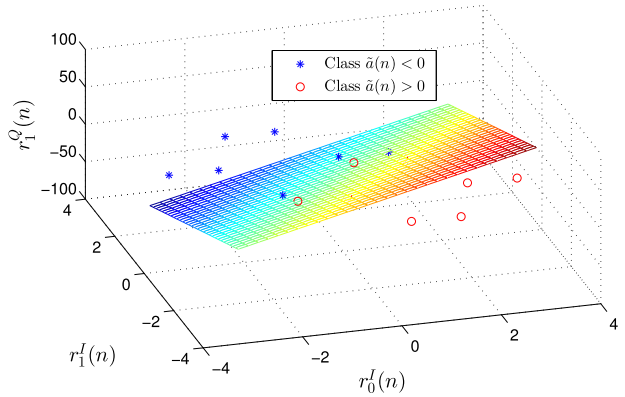


FIGURE 7. Hyperplane $r_1^Q(n) = 16.3r_0^I(n) - 15.8r_1^I(n)$ found by LB for channel 2.

when applied as an adaptive algorithm for smart antennas. In order to do that, the NB-Butterfly was created, which implements the same neural network structure present in the NE-Butterfly, but with the focus on spatial signal processing, instead of time-based.

A. ARCHITECTURE

The structure of the NB-Butterfly uses the similar structure presented in [18], illustrated in Figure 8, which presents an architecture composed of four MLP neural networks. Within that, two networks contribute to the processing of the signal in phase $\mathbf{r}^I(n)$, the MLP-II and MLP-IQ, whereas the others handle the processing of the signal in quadrature $\mathbf{r}^Q(n)$, being the MLP-QI and MLP-QQ. The signal $\mathbf{r}^I(n)$ is expressed as

$$\mathbf{r}^I(n) = \begin{bmatrix} r_0^I(n) \\ \vdots \\ r_p^I(n) \\ \vdots \\ r_{P-1}^I(n) \end{bmatrix} \tag{12}$$

and the signal $\mathbf{r}^Q(n)$ as

$$\mathbf{r}^Q(n) = \begin{bmatrix} r_0^Q(n) \\ \vdots \\ r_p^Q(n) \\ \vdots \\ r_{P-1}^Q(n) \end{bmatrix}. \tag{13}$$

These groups operate independently and in parallel from each other, as do the individual networks inside each group. The main objective of this arranging is so that each pair has input information from the real and complex part of the signal, with each MLP network combining them in a different manner leading to a better information acquisition from the various effects that may affect the channel, such as distortion, ISI and multi pathing.

The detailed architecture of the MLP-II, MLP-IQ, MLP-QI and MLP-QQ networks is similar to the structure presented in Figure 9. In each network, the hidden layer activation function $\varphi(\cdot)$ is the hyperbolic tangent function and the output layer function $\phi(\cdot)$ is the linear function.

Thus, the estimated phase signal in the output for each of the respective neural network group in the NB-Butterfly can be written as

$$\tilde{a}^I(n) = \tilde{a}^{II}(n) + \tilde{a}^{IQ}(n) \tag{14}$$

where

$$\tilde{a}^{II}(n) = \sum_{i=0}^{K-1} \left(w_i^1(n)\right)^{II} \tanh\left(\sum_{p=0}^{P-1} \left(w_{ip}^0(n)\right)^{II} r_p^I(n)\right), \tag{15}$$

$$\tilde{a}^{IQ}(n) = \sum_{i=0}^{K-1} \left(w_i^1(n)\right)^{IQ} \tanh\left(\sum_{p=0}^{P-1} \left(w_{ip}^0(n)\right)^{IQ} r_p^Q(n)\right) \tag{16}$$

and the estimated quadrature signal in the output for its neural network group is

$$\tilde{a}^Q(n) = \tilde{a}^{QI}(n) + \tilde{a}^{QQ}(n) \tag{17}$$

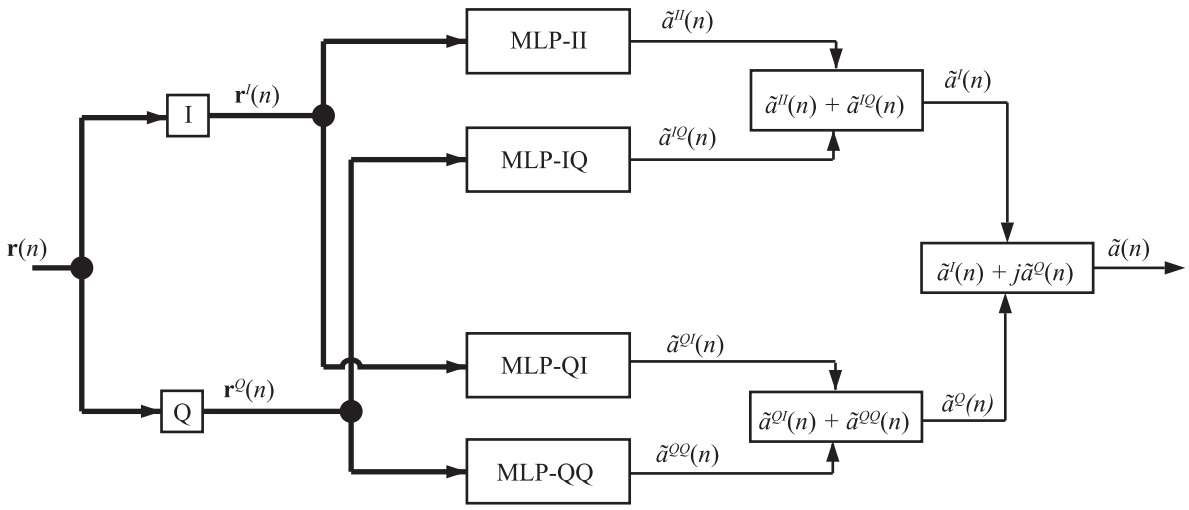


FIGURE 8. NB-Butterfly model.

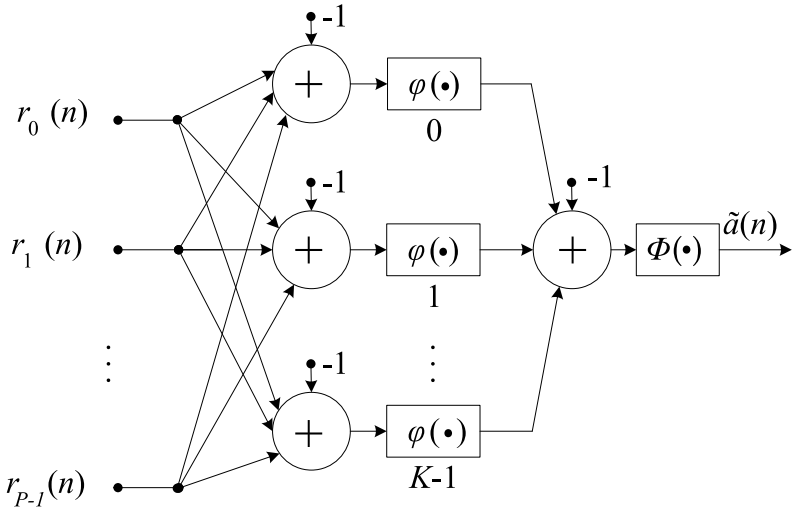


FIGURE 9. Neural beamformer architecture.

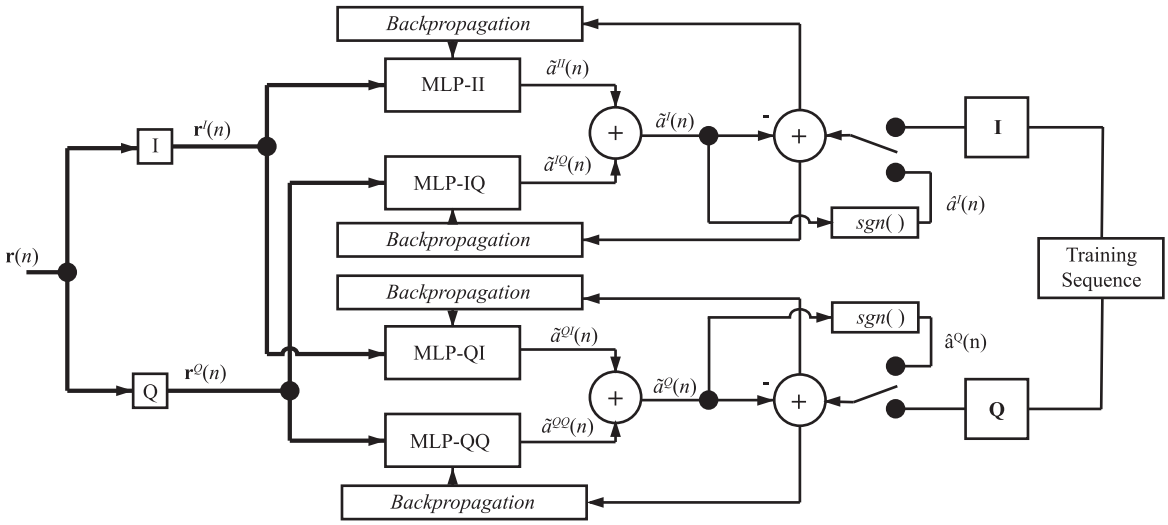


FIGURE 10. NE-butterfly training scheme.

where

$$\tilde{a}^{QI}(n) = \sum_{i=0}^{K-1} \left(w_i^1(n)\right)^{QI} \tanh\left(\sum_{p=0}^{P-1} \left(w_{ip}^0(n)\right)^{QI} r_p^I(n)\right), \tag{18}$$

$$\tilde{a}^{QQ}(n) = \sum_{i=0}^{K-1} \left(w_i^1(n)\right)^{QQ} \tanh\left(\sum_{p=0}^{P-1} \left(w_{ip}^0(n)\right)^{QQ} r_p^Q(n)\right) \tag{19}$$

where $\tilde{a}^I(n)$ and $\tilde{a}^Q(n)$ are the signals estimates $a^I(n)$ and $a^Q(n)$ respectively.

With the objective of respecting the hidden layer hyperbolic tangent activation function, the signals $a^I(n)$ and $a^Q(n)$ are normalized, where the symbols must obey a minimum distance d_{min} such that

$$d_{min} = \frac{2}{\sqrt{M}-1}. \tag{20}$$

B. TRAINING SCHEME

The first training scheme proposed for the NB-Butterfly is shown in Figure 10. It has two modes of training, supervised and non supervised. In the supervised mode, the error signal in phase and quadrature are given by

$$e^I(n) = a_{ref}^I(n) - \tilde{a}^I(n) \tag{21}$$

and

$$e^Q(n) = a_{ref}^Q(n) - \tilde{a}^Q(n), \tag{22}$$

where $a_{ref}^I(n)$ and $a_{ref}^Q(n)$ are the known training sequence. In the non supervised mode, the decision directed algorithm is

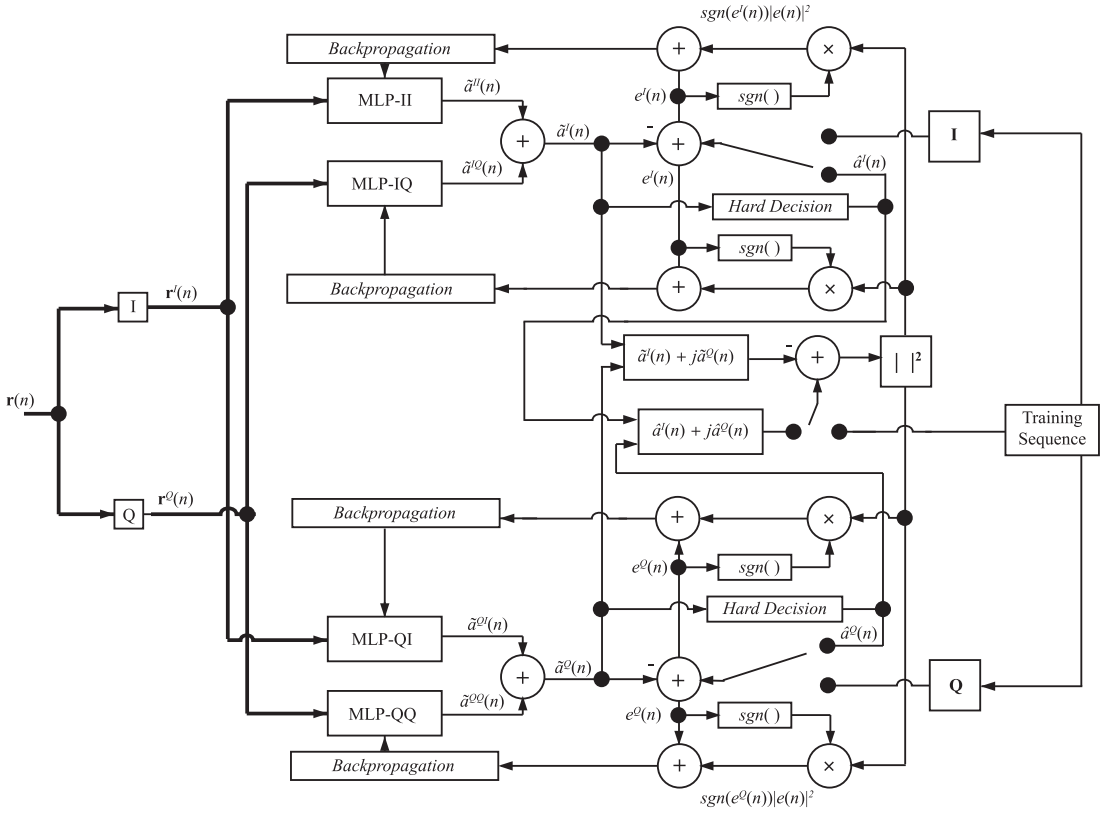


FIGURE 11. NE-butterfly-JE training scheme.

used, with signals in error and quadrature being respectively

$$e^I(n) = \hat{a}^I(n) - \tilde{a}^I(n) \tag{23}$$

and

$$e^Q(n) = \hat{a}^Q(n) - \tilde{a}^Q(n). \tag{24}$$

where $\hat{a}^I(n)$ and $\hat{a}^Q(n)$ are decision estimates of the received signal.

In order to achieve faster convergence of the NB-Butterfly, the new training scheme shown in Figure 11, named NB-Butterfly with joint errors (NB-Butterfly-JE), has combined errors fed back to the neural networks, which are described by

$$e^I_{JE}(n) = e^I(n) + sgn(e^I(n)) * |e(n)|^2 \tag{25}$$

and

$$e^Q_{JE}(n) = e^Q(n) + sgn(e^Q(n)) * |e(n)|^2, \tag{26}$$

where, in the non supervised mode, the decision directed algorithm is used, with the error $e(n)$ being given by

$$e(n) = \hat{a}(n) - \tilde{a}(n) \tag{27}$$

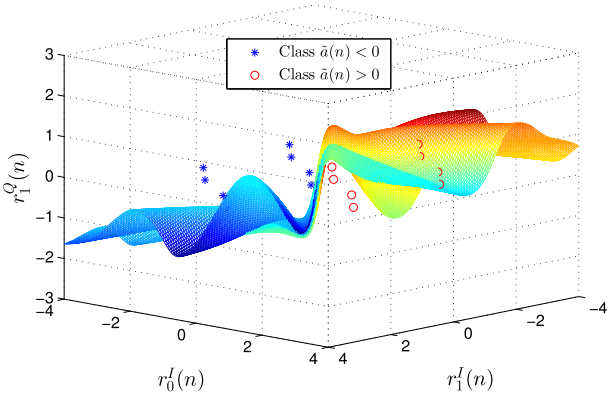


FIGURE 12. Non-linear classification found by NB-butterfly for channel 1.

while in the supervised mode it is given by

$$e(n) = a_{ref}(n) - \tilde{a}(n) \tag{28}$$

where

$$a_{ref}(n) = a^I_{ref}(n) + ja^Q_{ref}(n). \tag{29}$$

Figures 12 and 13 shows the non-linear classification associated with NB-Butterfly for channel 1 and channel 2 with

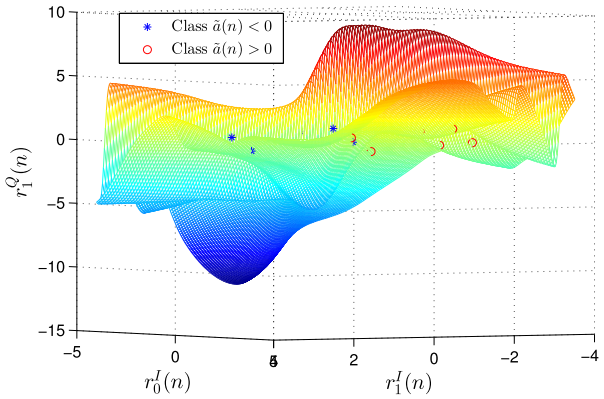


FIGURE 13. Non-linear classification found by NB-butterfly for channel 2.

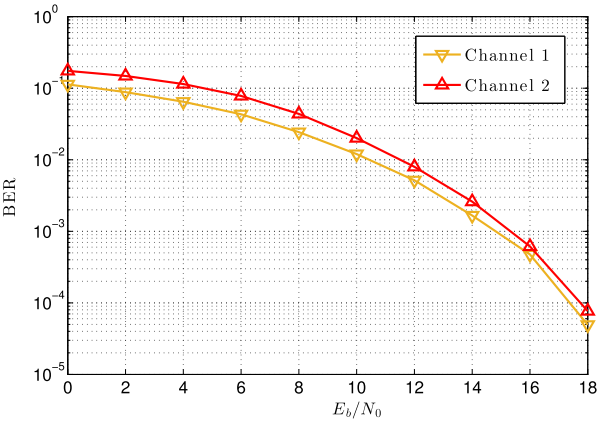


FIGURE 14. BER curve for the channels presented in Table 1. Results were obtained for a communication system (see Figure 1) using BPSK modulation with 2 antenna elements.

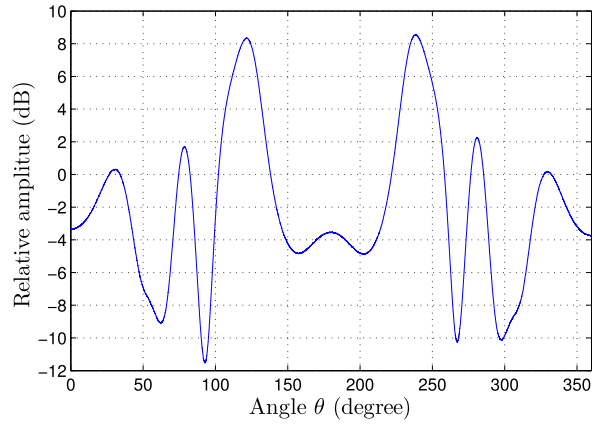


FIGURE 15. NB-butterfly radiation pattern for channel 1. BPSK system with $P = 2$ antennas elements.

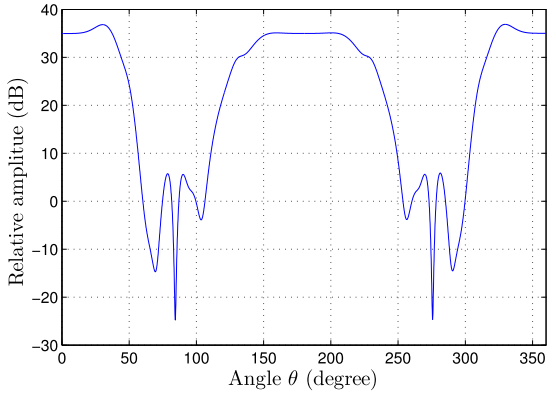


FIGURE 16. NB-butterfly radiation pattern for channel 2. BPSK system with $P = 2$ antennas elements.

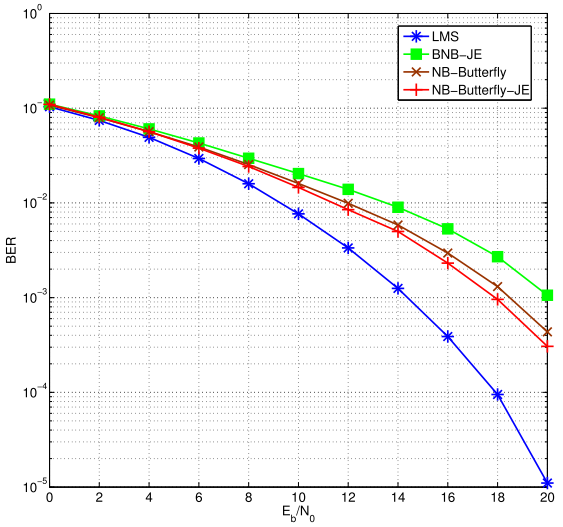


FIGURE 17. Performance curve of BER as a function of E_b/N_0 for the 4-QAM system (without channel coding) using the channel model for channel 3.

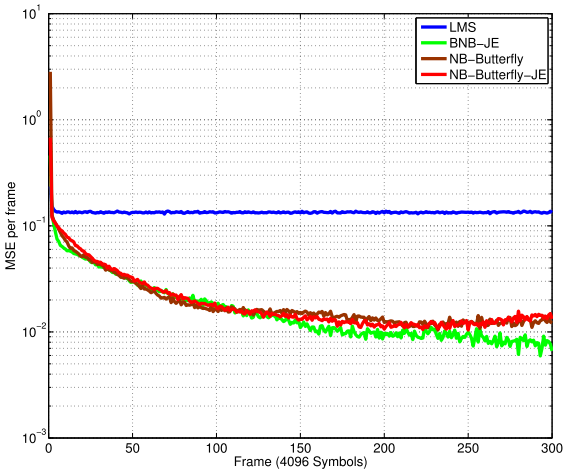
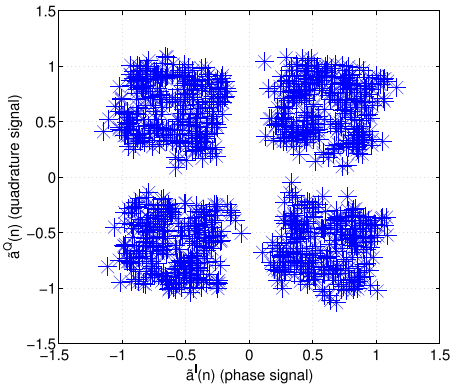


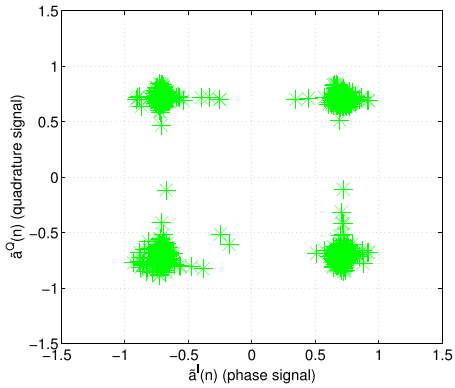
FIGURE 18. MSE curve for the 4-QAM system (without channel encoding) using the channel model for channel 3, with $E_b/N_0 = 20$dB.

$P = 2$ and $M = 2$ (BPSK system). The BER curves for this case are shown in Figure 14 and the radiation pattern curves are presented in Figures 15 and 16.

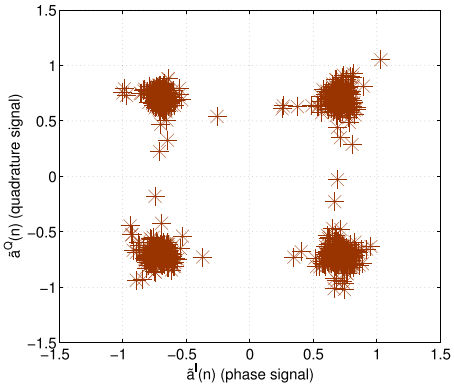
The Butterfly Neural Beamformer (NB-Butterfly) has distinct differences from the Neural Butterfly Equalizer (NE-Butterfly). In the equalization problem presented in [18], the nonlinearities are associated with the optical channel system and the neural butterfly equalizer proposed



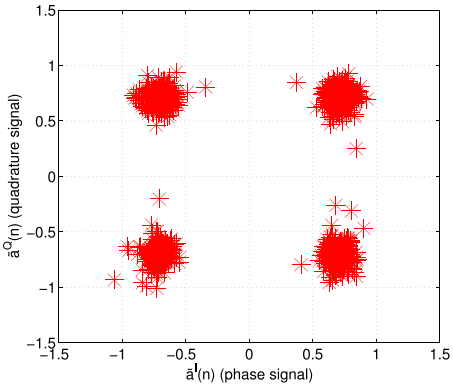
(a) LMS channel 3 constellation.



(b) BNB-JE channel 3 constellation.



(c) NB-Butterfly channel 3 constellation.



(d) NB-Butterfly-JE channel 3 constellation.

FIGURE 19. **Constellations for channel 3 with $E_b/N_0 = 20$dB.**

minimizes the problem regards to the complex taps channels. Already, this work shows that the beamforming nonlinearities appear when the number of antennas (P) is lower than the number of taps channels (L) ($P < L$) and differently optical channel with NE-Butterfly, the received signals always are multiplied by complex components (see Equation 2). Another difference is associated with the neural filter structure. The NE-Butterfly proposed in [18] focuses on processing a stream of N delayed signals while the NB-Butterfly works with a real-time stream of P signals from multiple antennas that arrive at different angles. Finally, another stark difference is that the NB-Butterfly-JE implements a new training method that significantly improves its performance, which is not implemented in the NE-Butterfly in [18].

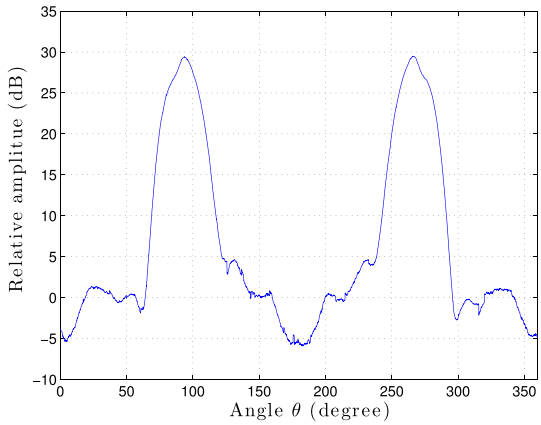


FIGURE 20. **NB-butterfly-JE radiation pattern for channel 3.**

IV. SIMULATIONS AND RESULTS

The first objective of the simulations is to analyze the performance of both the NB-Butterfly and NB-Butterfly-JE by evaluating their capacity to successfully perform the beamforming process on different channels. The second is to compare its performance with other proposals, where a LMS beamformer and a beamforming adapted BNB-JE were used to have both a linear and a neural beamformer to serve as a comparison basis. In order to evaluate the simulations, bit error rate (BER) curves in function of E_b/N_0 were traced,

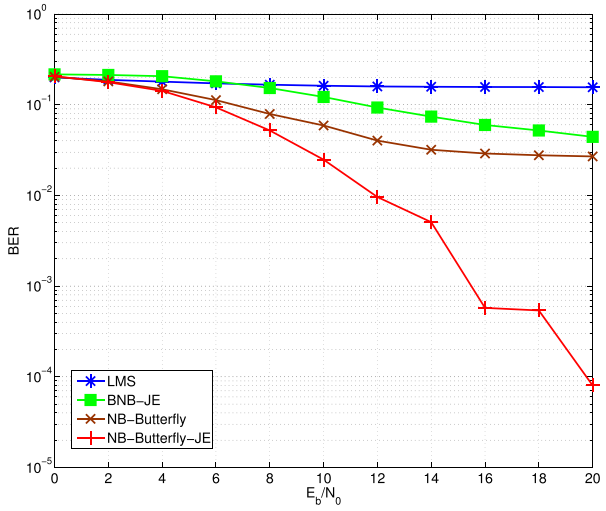


FIGURE 21. Performance curve of BER as a function of E_b/N_0 for the 4-QAM system (without channel coding) using the channel model for channel 4.

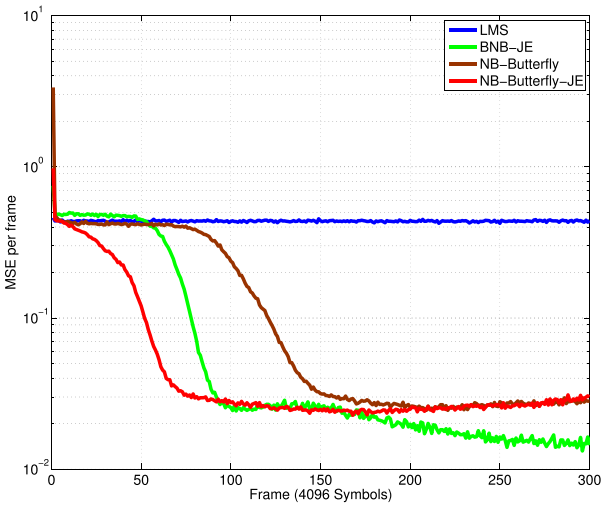
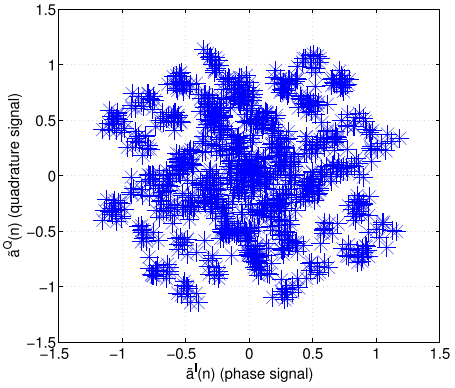
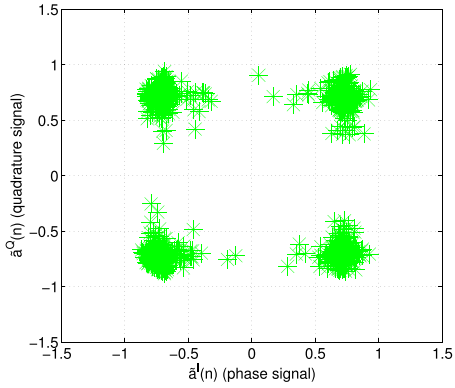


FIGURE 22. MSE curve for the 4-QAM system (without channel encoding) using the channel model for channel 4, with $E_b/N_0 = 20$dB.

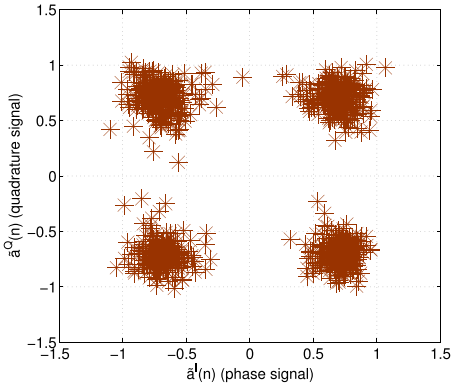
which denominates the relation between bit energy and power spectral density. The mean square error (MSE) curves of the backpropagation training for a E_b/N_0 of 20 dB were also acquired in order analyze the difference in performance between the beamformers regarding the converging time in function of the quantity of frames transmitted, where each



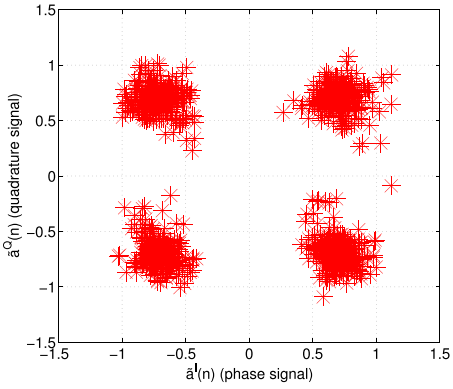
(a) LMS channel 4 constellation.



(b) BNB-JE channel 4 constellation.



(c) NB-Butterfly channel 4 constellation.



(d) NB-Butterfly-JE channel 4 constellation.

FIGURE 23. Constellations for channel 4 with $E_b/N_0 = 20$dB.

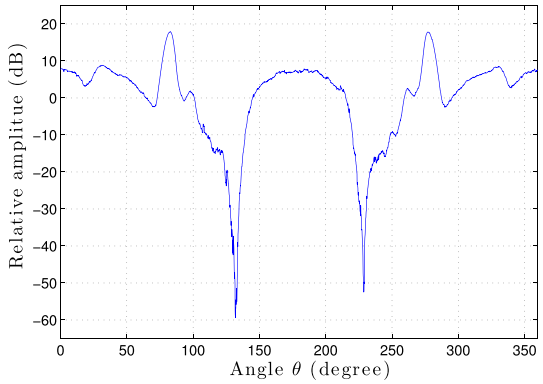


FIGURE 24. NB-butterfly-JE radiation pattern for channel 4.

frame contains 4096 symbols. Simulations were made for a 4-QAM digital communication system ($M = 4$), with two antennas.

Regarding channel 3, its BER curves are presented in Figure 17, its MSE curves are presented in Figure 18, the constellations for $E_b/N_0 = 20$dB are presented in Figure 19 and the radiation pattern for the NB-Butterfly-JE can be seen in Figure 20. This channel presents a linearly separable configuration and it is possible to infer that by looking at the LMS BER performance, as it is a fast and reliable beamformer for these cases, where from 4 dB onward, its speed of convergence cannot be matched by the NABs, however, when looking at the MSE curves and constellation points, it is clear that it is outperformed in that aspect, with the NABs managing to achieve lower MSE values and with a better clustering of the points, being on par in performance between them, with a small edge achieved by the BNB-JE in those regards.

For channel 4, its BER curves are presented in Figure 21, its MSE curves are presented in Figure 22, the constellations for $E_b/N_0 = 20$dB are presented in Figure 23 and the radiation pattern for the NB-Butterfly-JE can be seen in Figure 24. This channel's objective is to simulate a slightly non linearly separable configuration for two antennas, where, with the added complexity given by the complex gains, makes it also very hard for most neural networks to handle. By inspecting the BER curves, the NB-Butterfly-JE holds a great performance advantage over even the NB-Butterfly, implying that the new training method improves the search for the global minimum drastically, with its 12 dB performance never being outmatched by the other beamformers. Through an analysis of the MSE curves however, we see that the values between the NABs are quite close, but the convergence rate of the NB-Butterfly-JE is so fast in comparison to the others that it definitely helps in the lower BER numbers overall. Just like in channel 3, the constellations of the NABs are very similar, however in this case, the LMS wasn't able to properly achieve the correct point clustering for the 4-QAM.

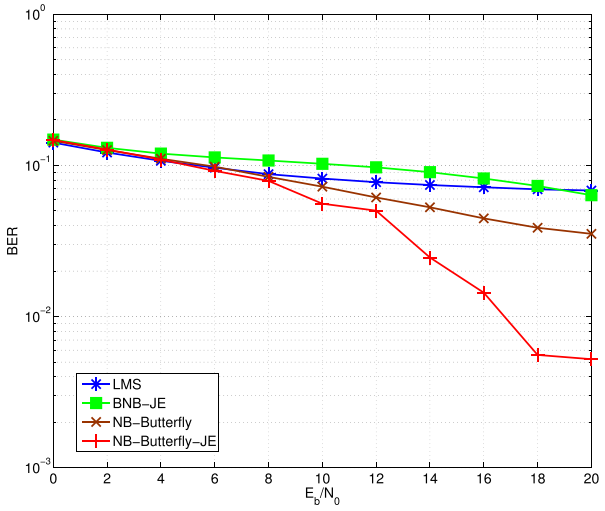


FIGURE 25. Performance curve of BER as a function of E_b/N_0 for the 4-QAM system (without channel coding) using the channel model for channel 5.

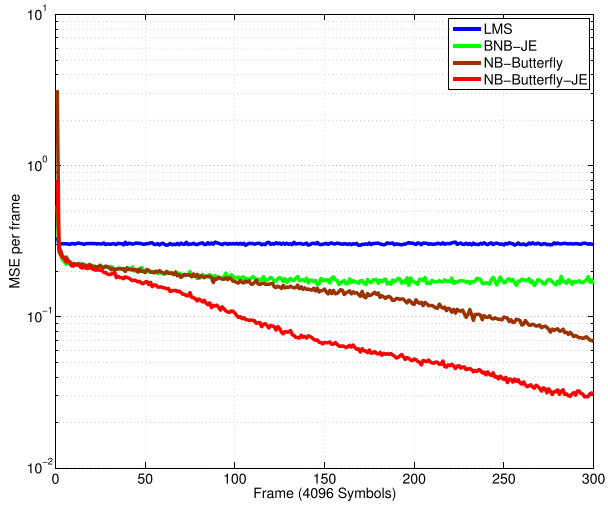
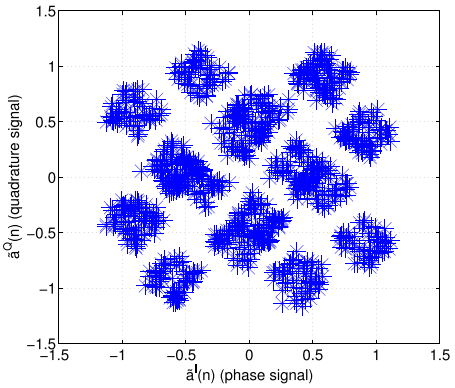
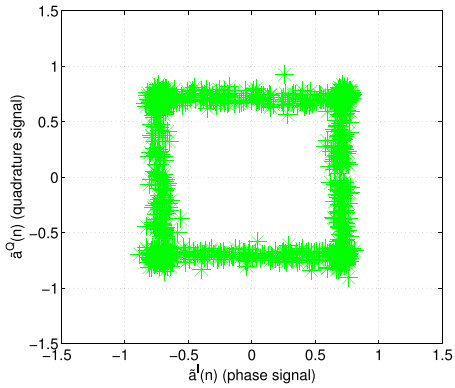


FIGURE 26. MSE curve for the 4-QAM system (without channel encoding) using the channel model for channel 5, with $E_b/N_0 = 20$dB.

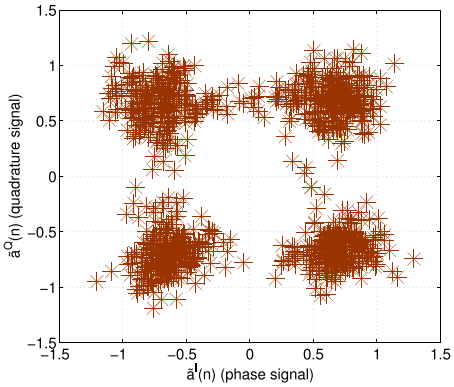
The BER curves for channel 5 are presented in Figure 25, its MSE curves are presented in Figure 26, the constellations for $E_b/N_0 = 20$dB are presented in Figure 27 and the radiation pattern for the NB-Butterfly-JE can be seen in Figure 28. This channel simulates a highly non linearly separable channel, also having complex gains makes it hard to handle for most neural networks that don't have a complex activation function or some strategy to deal with it. By looking at the BER curves, the NB-Butterfly-JE once again shows the best performance, from around 13 dB onwards its performance isn't reached by any other beamformers in this comparison, with the other neural networks barely making any progress at 20 dB. The analysis of the MSE curves and constellation points also supports the BER curves data, where the NB-Butterfly-JE has the best gradient descent,



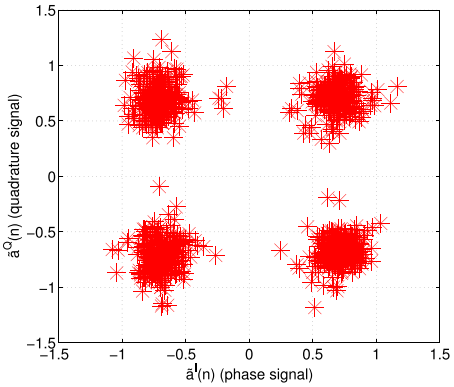
(a) LMS channel 5 constellation.



(b) BNB-JE channel 5 constellation.



(c) NB-Butterfly channel 5 constellation.



(d) NB-Butterfly-JE channel 5 constellation.

FIGURE 27. Constellations for channel 5 with $E_b/N_0 = 20$dB.

followed by the NB-Butterfly, once again supporting the benefits that came with the new training strategy. On the other hand, due to the complex nature of the channel, the BNB-JE wasn't able to achieve a good performance, being marginally better than the LMS on this case, but without managing to result in a clear clustering of the 4-QAM points on the constellation.

Figure 29 contains the BER curves for channel 6, Figure 30 presents its MSE curves, the constellations for $E_b/N_0 = 20$dB are presented in Figure 31 and the radiation pattern for the NB-Butterfly-JE can be seen in Figure 32. This channel simulates a non linearly separable channel with no complex gains, the objective is to show that a complex gain isn't needed in order to generate this kind of channel when looking at a beamformer working with a pair of antennas. In the analysis of the BER curves, by not reaching a low enough performance with 20 dB it becomes clear that the beamforming of this channel is difficult for all evaluated beamformers, with the NB-Butterfly-JE having the overall best performance, followed closely by the other NABs, with the LMS not managing to handle the nonlinearities. The MSE curves and constellation points also corroborate with these results, with

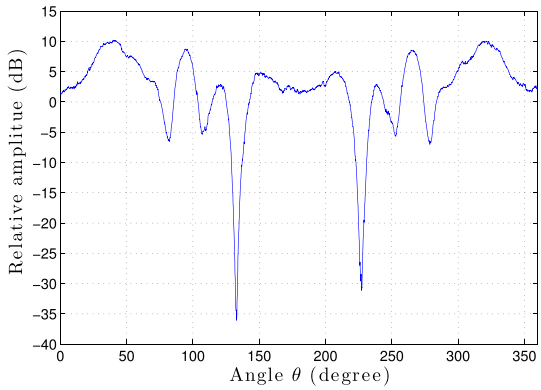


FIGURE 28. NB-butterfly-JE radiation pattern for channel 5.

the NABs all having similar curves and constellation points, while the LMS isn't able to reach the same error values nor make the correct clustering of the 4-QAM constellation points.

The results obtained in this work indicate that the NB-Butterfly is a viable option as a NAB architecture,

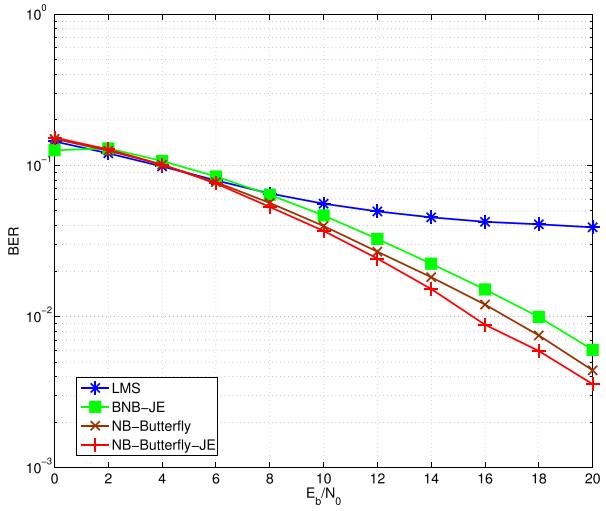


FIGURE 29. Performance curve of BER as a function of E_b/N_0 for the 4-QAM system (without channel coding) using the channel model for channel 6.

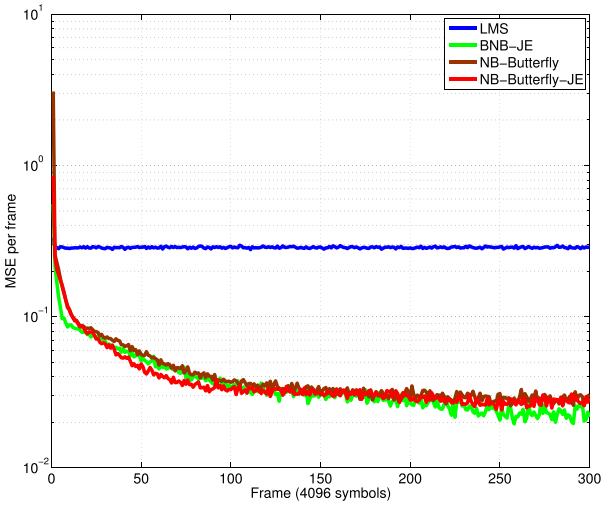
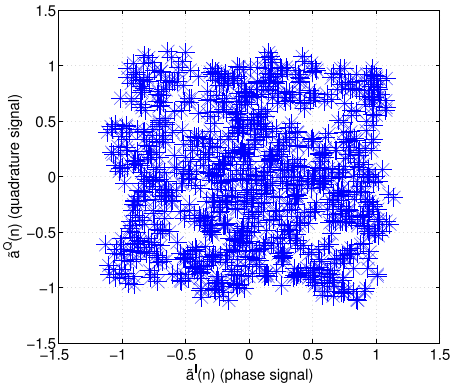
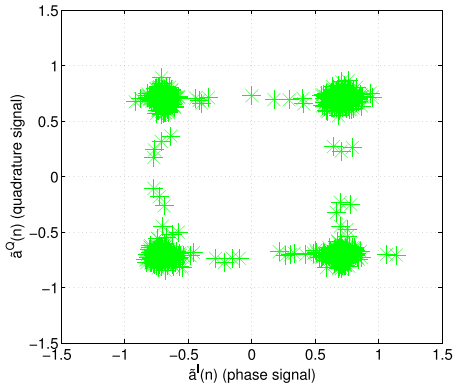


FIGURE 30. MSE curve for the 4-QAM system (without channel encoding) using the channel model for channel 6, with $E_b/N_0 = 20$dB.

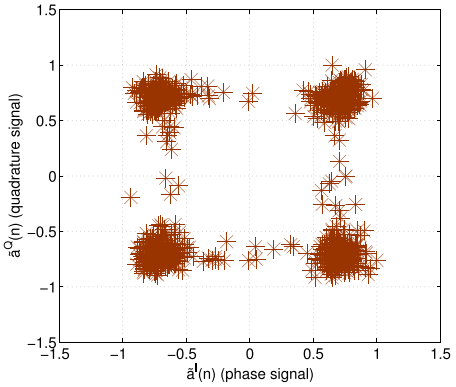
even more so when using the NB-Butterfly-JE's training method, that greatly speeds the convergence of the back-propagation algorithm, leading to a better performance, specially in the more complex situations, proving to be at least as viable as the other compared beamformers otherwise.



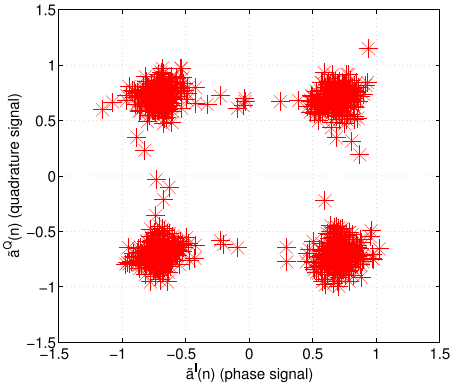
(a) LMS channel 6 constellation.



(b) BNB-JE channel 6 constellation.



(c) NB-Butterfly channel 6 constellation.



(d) NB-Butterfly-JE channel 6 constellation.

FIGURE 31. Constellations for channel 6 with $E_b/N_0 = 20$dB.

TABLE 3. Simulated channels with 4-QAM modulator.

Channel 3	Path gains	$\alpha_0 = 1.0$	$\alpha_1 = 0.5$	$\alpha_2 = 0.3$	$\alpha_3 = 0.1$
	Path delay	$\tau_0 = 0$	$\tau_1 = 3$	$\tau_2 = 6$	$\tau_3 = 9$
	AOA	$\theta_0 = 0°$	$\theta_1 = 30°$	$\theta_2 = 60°$	$\theta_3 = 90°$
Channel 4	Path gains	$\alpha_0 = 1.0$	$\alpha_1 = 0.7 + 0.7j$	$\alpha_2 = 0.3 + 0.3j$	$\alpha_3 = 0.1$
	Path delay	$\tau_0 = 0$	$\tau_1 = 3$	$\tau_2 = 6$	$\tau_3 = 9$
	AOA	$\theta_0 = 0°$	$\theta_1 = 30°$	$\theta_2 = 90°$	$\theta_3 = 100°$
Channel 5	Path gains	$\alpha_0 = 1.0$	$\alpha_1 = 0.5 + 0.5j$	$\alpha_2 = 0.3 + 0.3j$	$\alpha_3 = 0.1$
	Path delay	$\tau_0 = 0$	$\tau_1 = 3$	$\tau_2 = 6$	$\tau_3 = 9$
	AOA	$\theta_0 = 0°$	$\theta_1 = 30°$	$\theta_2 = 90°$	$\theta_3 = 100°$
Channel 6	Path gains	$\alpha_0 = 1.0$	$\alpha_1 = 0.7$	$\alpha_2 = 0.3$	$\alpha_3 = 0.1$
	Path delay	$\tau_0 = 0$	$\tau_1 = 3$	$\tau_2 = 6$	$\tau_3 = 9$
	AOA	$\theta_0 = 0°$	$\theta_1 = 30°$	$\theta_2 = 90°$	$\theta_3 = 130°$

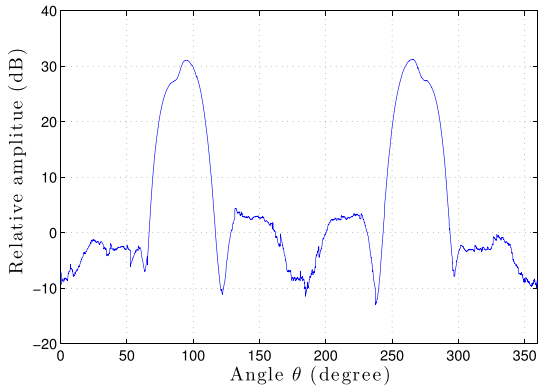


FIGURE 32. NB-butterfly-JE radiation pattern for channel 6.

TABLE 4. Parameters used in the simulated adaptive beamformers.

Structure	Number of inputs (P)	Neurons (hidden layer) (K)	Adaptation step (μ)	Momentum (α)
BNB-JE	2	32	0.001	10^{-6}
NB-Butterfly	2	32	0.001	10^{-6}
NB-Butterfly-JE	2	32	0.001	10^{-6}

V. CONCLUSIONS

This paper presented a new strategy based in ANN for the implementation of an adaptive beamformer, called NB-Butterfly, inspired on an adaptive equalizer structure, the NE-Butterfly, there was also a new training strategy proposed for this NAB, being the NB-Butterfly-JE. As the name implies, four MLP neural networks are used, MLP-II, MLP-IQ, MLP-QQ and MLP-QQ, connected in an architecture that is similar to a butterfly with its open wings, thus the name. The objective of this beamformer is to perform the beamforming in phase and quadrature of a modulated signal that runs through a channel that has real or complex gains. In its training scheme, the error is calculated independently for each MLP pair and can be supervised or not for the NB-Butterfly, and for the NB-Butterfly-JE a heuristic is added where the individual errors from the four networks are combined with the error from the main signal generated by these networks in order to speed up the gradient descent of the backpropagation training function, thus leading to the name Joint Error.

The NB-Butterfly was used in a simulated channel which has the main problems found in wireless communication channels, which are the nonlinearities, different AOA and possible complex gains. Through simulations it was possible to test the beamformers using different values of E_b/N_0 and to compare them with other previously tested beamformers, such as the LMS and the BNB-JE, which was adapted from the BNE-MLP-BP-JE, where the NB-Butterfly proved effective in beamforming the majority of the channels that were evaluated, making it a valid proposal to use in adaptive beamforming.

REFERENCES

[1] S. S. Haykin, *Communication Systems*, 4th ed. Hoboken, NJ, USA: Wiley, 2001.
[2] J. C. Liberti and T. S. Rappaport, *Smart Antennas for Wireless Communications: IS-95 and Third Generation CDMA Applications*. Upper Saddle River, NJ, USA: Prentice-Hall, 1990.
[3] J. Proakis, *Digital Communications*. New York, NY, USA: McGraw-Hill, 2000.
[4] S. Chen, L. Hanzo, and A. Wolfgang, "Nonlinear multiantenna detection methods," *EURASIP J. Appl. Signal Process.*, vol. 9, pp. 1225–1237, Jan. 2004.
[5] S. Chen, A. Wolfgang, and L. Hanzo, "A robust nonlinear beamforming assisted receiver for BPSK signalling," in *Proc. IEEE Veh. Technol. Conf.*, Sep. 2005, vol. 3, no. 62, pp. 1921–1925.
[6] S. Chen, A. Wolfgang, C. J. Harris, and L. Hanzo, "Adaptive nonlinear least bit error-rate detection for symmetrical rbf beamforming," *Neural Netw.*, vol. 21, pp. 358–367, Mar./Apr. 2008.
[7] S. Chen, L. Hanzo, and S. Tan, "Symmetric complex-valued RBF receiver for multiple-antenna-aided wireless systems," *IEEE Trans. Neural Netw.*, vol. 19, no. 9, pp. 1659–1665, Sep. 2008.
[8] N. Xie and Y. Zhou, "A new nonlinear beamformer for CDMA wireless communications," in *Proc. IEEE Int. Conf. Netw., Sens. Control*, Apr. 2006, pp. 486–491.
[9] S. Chen, L. Hanzo, and A. Wolfgang, "Kernel-based nonlinear beamforming construction using orthogonal forward selection with the Fisher ratio class separability measure," *IEEE Signal Process. Lett.*, vol. 11, no. 5, pp. 478–481, May 2004.
[10] D. Comminiello, M. Scarpiniti, L. A. Azpicueta-Ruiz, J. Arenas-García, and A. Uncini, "Combined nonlinear filtering architectures involving sparse functional link adaptive filters," *Signal Process.*, vol. 135, pp. 168–178, Jun. 2017.
[11] S. Shaikh and D. K. Panda, "Linear, non-linear adaptive beamforming algorithm for smart antenna system," in *Proc. IEEE Int. Conf. Comput., Commun. Control*, Sep. 2015, pp. 1–4.
[12] A. Rawat, R. N. Yadav, and S. C. Shrivastava, "Neural network applications in smart antenna arrays: A review," *AEU-Int. J. Electron. Commun.*, vol. 66, no. 11, pp. 903–912, Nov. 2012.
[13] K.-L. Du, A. K. Y. Lai, K. K. M. Cheng, and M. N. S. Swamy, "Neural methods for antenna array signal processing: A review," *Signal Process.*, vol. 82, pp. 547–561, Apr. 2002.

[14] Z. D. Zaharis, K. A. Gotsis, and J. N. Sahalos, "Comparative study of neural network training applied to adaptive beamforming of antenna arrays," *Progr. Electromagn. Res.*, vol. 126, pp. 269–283, Mar. 2012.
[15] R. Savitha, S. Vigneswaran, S. Suresh, and N. Sundararajan, "Adaptive beamforming using complex-valued radial basis function neural networks," in *Proc. TENCON*, Jan. 2009, pp. 1–6.
[16] A. H. E. Zooghby, C. G. Christodoulou, and M. Georgiopoulos, "Neural network-based adaptive beamforming for one- and two-dimensional antenna arrays," *IEEE Trans. Antennas Propag.*, vol. 46, no. 12, pp. 1891–1893, Dec. 1998.
[17] B. Li, T. N. Sainath, R. J. Weiss, K. W. Wilson, and M. Bacchiani, "Neural network adaptive beamforming for robust multichannel speech recognition," in *Proc. INTERSPEECH*, 2016, pp. 1976–1980.
[18] T. F. B. de Sousa and M. A. C. Fernandes, "Butterfly neural equalizer applied to optical communication systems with two-dimensional digital modulation," *Opt. Express*, vol. 26, no. 23, pp. 30837–30850, Nov. 2018.
[19] M. A. C. Fernandes, "Neural equalization applied to systems with bidimensional digital modulation," *Neural Comput. Appl.*, vol. 25, nos. 7–8, pp. 2057–2066, Dec. 2014.
[20] T. F. B. de Sousa and M. A. C. Fernandes, "Bi-dimensional neural equalizer applied to optical receiver," in *Proc. BRICS Congr. Comput. Intell. 11th Brazilian Congr. Comput. Intell.*, Sep. 2013, pp. 34–39.
[21] T. F. B. de Sousa and M. A. C. Fernandes, "Multilayer perceptron equalizer for optical communication systems," in *SBMO/IEEE MTT-S Int. Microw. Symp. Dig. Optoelectron. Conf. (IMOC)*, Aug. 2013, pp. 1–5.
[22] T. F. B. de Sousa and M. A. C. Fernandes, "Multilayer perceptron equalizer for optical communication systems," *WSEAS Trans. Commun.*, vol. 13, no. 1, pp. 462–469, 2014.
[23] T. F. B. de Sousa, D. S. Arantes, and M. A. C. Fernandes, "Adaptive beamforming applied to OFDM systems," *MDPI Sensors*, vol. 18, no. 10, p. 3558, 2018.

TIAGO F. B. DE SOUSA was born in Teresina, Brazil. He received the B.S. degree in computer engineering and the M.S. degree in electrical and computer engineering from the Federal University of Rio Grande do Norte, Natal, Brazil, in 2011 and 2014, respectively. He is currently a Game Development Professor with the Federal Institute of Rio Grande do Norte, Natal. His research interests include artificial intelligence and digital signal processing.

MARCELO A. C. FERNANDES was born in Natal, Brazil. He received the B.S. degree in electrical engineering, the M.S. degree in electrical engineering, from the Federal University of Rio Grande do Norte, Natal, in 1997 and 1999, respectively, and the Ph.D. degree in electrical engineering from the University of Campinas, Campinas, Brazil, in 2010. From 2015 to 2016, he was a Visiting Researcher with the Centre Telecommunication Research (CTR), King's College London, London, U.K. He is currently an Adjunct Professor with the Department of Computer Engineering and Automation, Federal University of Rio Grande do Norte, Natal. He is also the Leader of the Research Group on Embedded Systems and Reconfigurable Computing (RESRC) and a Coordinator with the Laboratory of Machine Learning and Intelligent System (LMLIS). His research interests include artificial intelligence, digital signal processing, embedded systems, reconfigurable hardware, and the tactile internet. He has authored and co-authored many scientific papers and practical studies with reconfigurable computing on FPGA to accelerate artificial intelligence algorithms.