NEW BUSSGANG METHODS FOR BLIND EQUALIZATION

João B. Destro Filho[1,2], Gérard Favier[1] and João M. Travassos Romano[3]

1 Laboratoire I3S - CNRS/UNSA
250 Av. Albert Einstein, Bât. 4
Sophia Antipolis
06560 VALBONNE - FRANCE
2 Scholar of CNPq - Brasília - Brazil

3 DECOM - FEEC - UNICAMP
University of Campinas
Caixa Postal 6101
13081-970 Campinas - SP BRAZIL

Abstract: This paper analyses Bussgang blind algorithms [1]-[5] in terms of the minimization of an associated cost function. Then, two new blind algorithms are proposed and evaluated by computer simulation.

1. INTRODUCTION

Traditional Bussgang techniques [1]-[5] are the oldest methods for channel blind equalization. These methods can be interpreted as stochastic gradient algorithms minimizing a nonlinear cost function [6].

Figure 1 illustrates a data transmission system using a blind equalizer.

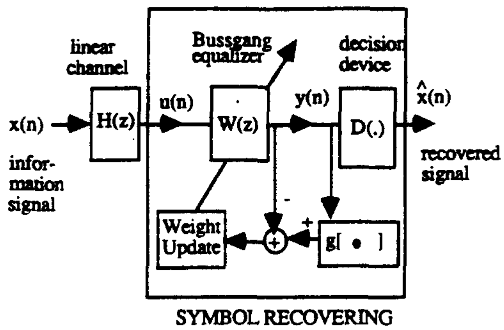

Figure 1 - Data transmission system

Bussgang algorithms can be written with the following equations:

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \mu(n)\, e(n)\, \mathbf{U}(n) \quad (1)$$
$$e(n) = g_B[y(n)] - y(n) \quad (2)$$
$$y(n) = \mathbf{U}^T(n)\, \mathbf{W}(n) \quad (3)$$

where:

$\mathbf{W}(n)=[w_0(n)....w_{L-1}(n)]^T$ is the equalizer weight vector, at time n, $w_i(n)$ is the i-th equalizer coefficient at time n, L is the equalizer length;
$\mu(n)$ is the step-size of the algorithm, at time n;
$g_B[.]$ is a memoryless nonlinear function;
$\mathbf{U}(n)=[u(n)\ u(n-1)\ ...\ u(n-L+1)]^T$ is the vector of L sample measurements received at the channel output, and y(n) is the equalizer output.

So, each blind equalization algorithm is defined by a special choice of $g_B[.]$. When this nonlinearity is chosen as the decision device D(.), we get the well-known decision-directed (DD) algorithm.
The DD algorithm is often used in association with another blind algorithm (for example, see [8]-[9]), due to its simplicity and its good convergence properties once the eye pattern is open [7]. That leads to dual-mode algorithms. However, when the channel eye is closed, DD adaptation should be avoided, or at least improved. In [9], two solutions are reported to achieve such improvement: step-size normalization /adaptation, and application of a big step-size.
From an analysis of [1]-[5], it results that two main approaches can be used for deriving Bussgang algorithms. In the first approach [4]-[5], the algorithms are obtained by minimizing a cost function, as it is usually the case for establishing adaptive algorithms. In the second approach [1]-[3], the equations of the algorithms are a priori chosen under the form (1)-(3) with a special choice of $g_B[.]$, and a possible associated cost function is then discussed.
This paper is mainly concerned with the second approach. In section 2, we define a general cost function and derive the "exact" stochastic gradient algorithm resulting from the minimization of this cost function. The corresponding equations are slightly different from equations (1)-(2). We give the conditions under which our algorithm is

identical to the algorithms of [4]-[5]. In sections 3 and 4, we discuss how DD/Sato's algorithm [1]-[2] and Bellini's algorithm [3] are linked with our general algorithm. Then, we propose a modified DD algorithm and a modified Bellini's algorithm. Finally, some simulation results are presented in section 5 to illustrate the behaviour of the proposed algorithms.

2. A GENERAL COST FUNCTION FOR BUSSGANG ALGORITHMS

In this paper, we assume that the transmitted signal x(n) is a PAM, iid, non-gaussian, zero-mean signal, H(z) is a linear model, and $g_B[.]$ is a differentiable function, or at least a piecewise-differentiable function. Under this last hypothesis, $g_B[.]$ will be written g[.].

Consider the following general cost function:

$$J = \frac{1}{2} E\left\{\left(g[y(n)] - y(n)\right)^2\right\} \qquad (4)$$

where E[.] represents the mathematical expectation operator.

The stochastic gradient algorithm resulting from the minimization of the criterion (4) with respect to the parameters vector **W(n)**, is given by the general equations (5)-(6):

$$\Delta \mathbf{W(n)} = \mathbf{W(n+1)} - \mathbf{W(n)} =$$

$$= \mu f[y(n)]\left(g[y(n)] - y(n)\right)\mathbf{U(n)} \qquad (5)$$

$$f[y(n)] = 1 - \frac{\partial g[y(n)]}{\partial y(n)} \qquad (6)$$

Notice that equation (5) is slightly different from the standard equations (1)-(2) due to the presence of the term f[y(n)]. In fact, a comparison of these equations suggests that the term $\mu f(y(n))$ in (5) could be viewed as a variable step-size.

The conditions under which the algorithms (1)-(2) and (5)-(6) are the same, are given below:

$$g_B(n) = g(y(n)) \qquad (7)$$

$$f[y(n)] = 1 - \frac{\partial g[y(n)]}{\partial y(n)} = 1$$

$$\Rightarrow \frac{\partial g[y(n)]}{\partial y(n)} = 0 \quad \forall y(n) \qquad (8)$$

Table I contains the expressions of g[.] associated with the Bussgang algorithms of [1]-[5].

Table I	
Algorithms	Associated g[y(n)]
DD [1]	D(.)
Sato [2]	S.sign[y(n)]
Bellini [3]	see eq. (18)
Godard (q = 2) [4]	$R_p - \lvert y(n)\rvert^p + y(n)$
Shalvi-Weinstein [5]	$y(n)^2 + y(n)$

where:

$$S = \frac{E\left[x(n)^2\right]}{E\left[\lvert x(n)\rvert\right]} \qquad (9)$$

$$\text{sign}[y(n)] = \begin{cases} +1 & \text{if } y(n) \geq 0 \\ -1 & \text{if } y(n) < 0 \end{cases} \qquad (10)$$

$$R_p = E\left[\lvert x(n)\rvert^{2p}\right] \Big/ E\left[\lvert x(n)\rvert^{p}\right] \qquad (11)$$

In the following, we demonstrate how equations (4)-(6) may represent the algorithms in [4]-[5].

For PAM signals, the Godard's cost function and algorithm are respectively given by the following equations [4]:

$$J_G = \frac{1}{2} E\left\{\left(R_p - \lvert y(n)\rvert^p\right)^2\right\} \qquad (12)$$

$$\mathbf{W(n+1)} = \mathbf{W(n)} +$$

$$\mu p\, y(n)\lvert y(n)\rvert^{p-2}\left(R_p - \lvert y(n)\rvert^p\right)\mathbf{U(n)} \qquad (13)$$

By replacing the Godard's g[.] defined in table I in (4)-(6), and recalling that y(n) = sign[y(n)] |y(n)|, one get equations (12)-(13).

For real PAM signals, the Shalvi-Weinstein's cost function [5] is given by:

$$J_s = (1/2)\, E\{y^4(n)\} \qquad (14)$$

$$\text{subject to } E\{y^2(n)\} = E\{x^2(n)\} \qquad (15)$$

Notice that Shalvi-Weinstein's algorithm

This minimization by means of the stochastic gradient algorithm leads to the following equation:

$$\mathbf{W}(n+1) = \mathbf{W}(n) + 2\mu\, y^3(n)\, \mathbf{U}(n) \qquad (16)$$

By replacing the Shalvi-Weinstein's g[.] as given in table I, in equations (4)-(6), one get equations (14)-(15).

In the two next sections, we analyse under which conditions the DD [1], Sato's [2] and Bellini's algorithms [3] satisfy equations (7)-(8). To our knowledge, this interpretation of the Bussgang algorithms [1]-[3] has never been addressed by the literature. Surprising as it may seem, it has always been stated, since the pioneering work of [2]-[3] and until recent work [6], [10]-[11], that equations (1)-(2) are associated to the minimization of the cost function (4), without mentioning the term f[y(n)].

3. A NEW MODIFIED DD ALGORITHM

For PAM signals, the decision device D(.) is generally a quantizer, so g[.] is not differentiable. Consider the simple case of a 2-PAM signal, where g[.] is given by (10). In order to enable our mathematical analysis, we approximate the sign function by an hyperbolic tangent h(y,β) with a parameter β that tends to infinity [12]:

$$\mathrm{sign}[y] \cong \lim_{\beta\to\infty} h(y,\beta) = \lim_{\beta\to\infty} \frac{\left(1-e^{-\beta y}\right)}{\left(1+e^{-\beta y}\right)} \qquad (17)$$

Figure 2 presents h(y,β) and its derivative for various values of β. Clearly, condition (8) is not satisfied for every y(n). In fact, y(n) is theoretically a zero-mean signal, so the derivative g'(y) will take large values almost all the time.
As a consequence, the term μf(y(n)) in (5) will also take large values. If we consider this term as a variable step-size, algorithm (5) could be compared to the traditional DD algorithm, which is equivalent to equation (1) when g[.] is the sign function. From this point of view, one may easily understand the improvement of DD adaptation [9] when the step-size is set to a big value.
Notice that all the above remarks also apply to the Sato's algorithm.

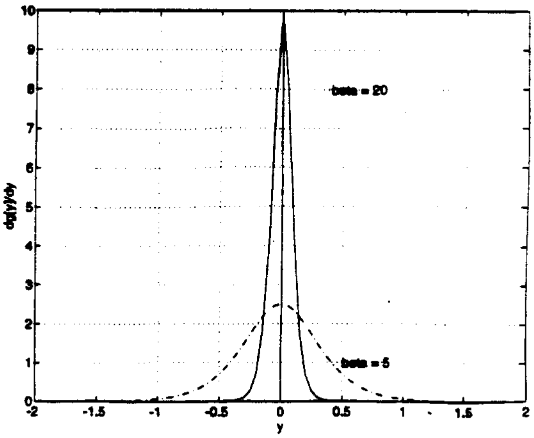

Figure 2 - Function h(y,β)

In this way, we propose a new modified DD algorithm, given by equations (5)-(6) with g[.] chosen as (17), β being set to a big value.

4. A NEW MODIFIED BELLINI'S ALGORITHM

Consider the simple case of a 2-PAM signal, and suppose a white gaussian convolutional noise of power σ^2 [10]. In this case, the Bellini's g(y) is given by the following expression:

$$g(y) = \frac{\exp\left(\frac{-(y-1)^2}{2\sigma^2}\right) - \exp\left(\frac{-(y+1)^2}{2\sigma^2}\right)}{\exp\left(\frac{-(y-1)^2}{2\sigma^2}\right) + \exp\left(\frac{-(y+1)^2}{2\sigma^2}\right)} \qquad (18)$$

This g(y) and its derivative are quite similar to those in figure 2. Clearly, (8) is not verified. We then propose a new Bellini's algorithm given by equations (5)-(6) with g(y) defined by (18).

5. SIMULATIONS

In this section, we present some simulation results to illustrate the behaviour of the two proposed algorithms and to compare their performance with that of the standard Bussgang algorithms [1],[3], in terms of convergence speed. A MA nonminimum phase channel given by $H(z) = 1 - 2.6z^{-1} + 1.2z^{-2}$ was simulated, with a signal-to-noise ratio of 40 dB. Monte-Carlo type simulations were carried out with 30 input data sequences x(n). Figures 4 and 5 present the simulation results.

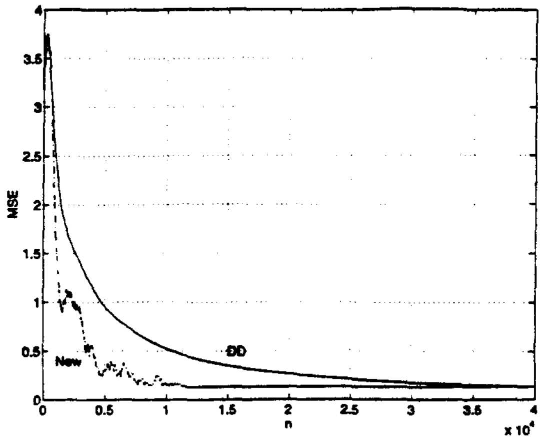

Figure 4 - Comparison of DD and modified DD algorithms

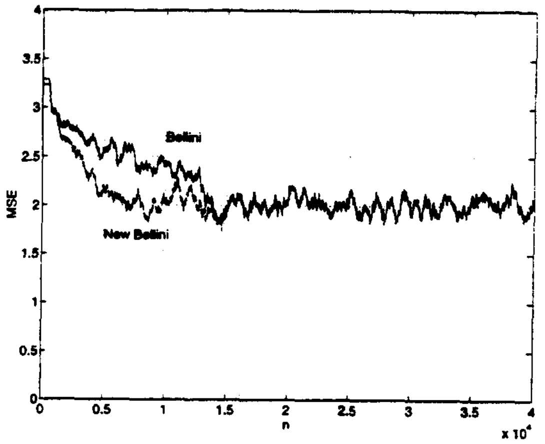

Figure 5 - Comparison of Bellini and modified Bellini' algorithms

From these simulation results, we can conclude that the proposed algorithms converge more fastly than the standard DD and Bellini algorithms. However, we have to notice that the modified DD algorithm presents some problems for big values of f(y) in the transitory period.

6. CONCLUSIONS

Bussgang algorithms for blind equalization have been analysed in this paper. An interpretation of these algorithms in terms of the minimization of a general cost function has been carrieed out. Two new blind algorithms (modified DD and modified Bellini's algorithms) have been proposed. Simulation results have been shown to illustrate the good behaviour of these new blind algorithms that converge faster than their standard counterparts.

7. ACKNOWLEDGEMENTS

We would like to thank CNPq and CPqD-Telebras for the finantial support of this work.

8. REFERENCES

[1] R. W. Lucky; The Bell Syst. Tech. J., vol. 45, pp. 255-286, Feb. 1966.

[2] Y. Sato; IEEE Trans. Commun., vol COM-23, pp. 679-682, June 1975.

[3] S. Bellini; Proc. of GLOBECOM'86, pp. 1634-1640, Houston, TX, 1986.

[4] D. N. Godard; IEEE Trans. Commun., vol. COM-28, no. 11, pp. 1867-1875, Nov. 1980.

[5] O. Shalvi and E. Weinstein; IEEE Trans. Inf. Theory, vol. IT-36, no. 2, pp. 312-321, March 1990.

[6] S. Haykin; "Adaptive filter theory", 2nd. ed., Prentice-Hall, Englewood Cliffs, NJ, 1991.

[7] O. Macchi and E. Eweda; IEEE Trans. Inf. Theory., vol. IT-30, no. 2, pp. 161-176, March 1984.

[8] A. Benveniste and M. Goursat; IEEE Trans. Commun., vol. COM-32, no. 8, pp. 871-883, Aug. 1984.

[9] K. Hilal; Technical note NT/PAB/RPE/ETP/3633, CNET, Nov. 1993.

[10] S. Bellini, in "Blind deconvolution", S. Haykin, editor, pp. 8-59, Prentice-Hall Inc., NJ, 1994.

[11] Z. Ding, C. R. Johnson, and R.A. Kennedy, in "Blind deconvolution", S. Haykin, editor, pp. 60-120, Prentice-Hall Inc., NJ, 1994.

[12] E. Corwin, A. Logar and W. Oldham, IEEE Trans. Neural Net., vol. 5, no. 3, 507-508, May 1994.