Blind Equalization and Carrier Recovery Using a "Stop-and-Go" Decision-Directed Algorithm

GIORGIO PICCHI AND GIANCARLO PRATI, MEMBER, IEEE

Abstract—**We show that the standard decision-directed estimated-gradient adaptation algorithm for joint MSE equalization and carrier recovery, normally utilized in the open-eye condition, can be turned into an algorithm providing effective blind convergence in the MSE sense, usable in the closed-eye startup phase with no need of a known training sequence. This is obtained by means of a simple flag telling both the equalizer and the synchronizer whether the current output error with respect to the decided symbol is sufficiently reliable to be used. If not, adaptation is stopped for the current iteration. In the paper, this "stop-and-go" decision-directed algorithm is presented for both linear and decision-feedback MSE complex equalizers with joint blind carrier recovery. Simulation results demonstrate the effectiveness of the proposed technique.**

I. INTRODUCTION

After the pioneering work by Sato [1], blind equalization for one- and two-dimensional communication systems has become an area of great interest for research and industry applications. As is known, blind equalizers converge without resorting to a known training sequence or to a condition of limited distortion. As a consequence, equalizer convergence does not require any preliminary carrier phase recovery. On the contrary, blind carrier phase tracking can be jointly performed at the equalizer output in a decision-aided mode, although during convergence, the decisions are mostly incorrect.

The usefulness of these properties has been pointed out by Godard [2] in the context of multipoint networks. Blind joint equalization and carrier recovery may also find application in digital radio links over multipath fading channels. In this case, an arising deep selective fade produces, when the channel distortion becomes too severe for the equalizer, a loss of the decision-aided carrier synchronism (lock-out) while transmission continues. Normally, the equalizer is thus reset, and the carrier is recovered (lock-in) and the equalizer is reactivated only when the unequalized channel gets better enough to allow for correct decisions. A blind joint equalizer and carrier recovery system would instead resume correct operation earlier, independently of the fact that the equalizer tap settings and the phase offsets initially cause most decisions on the data stream to be incorrect.

The aforementioned papers by Sato [1] and Godard [2], and the more recent paper by Benveniste and Goursat [3], approach the problem of blind equalization by introducing new nonconvex cost functions, different from the traditional mean-square error (MSE) functional used for the trained equalizer. Under weak conditions, these cost functions are shown to characterize the intersymbol interference sufficiently well, while their stochastic minimization can be performed by using ad hoc generated control signals with no knowledge of the transmitted data. A feature common to these algorithms is that performance after the initial convergence is not fully satisfactory, since the accuracy in locating the tap settings for minimum MSE (a parameter related to the measured output error) is poor. This is due either to the noncoincidence of the two minima, of the used cost function, and of the MSE, as in Godard's algorithm, or to the large residual variance of the error signal after the initial convergence, as compared to standard trained equalization, as in the Sato algorithm. Benveniste and Goursat propose an heuristic algorithm automatically switching from the blind Sato to the smoother conventional decision-directed (DD) algorithm when convergence is completed. Side by side with the above algorithms, the literature on the subject offers some interesting studies on the self-learning capabilities of the conventional DD algorithm when used from the beginning of the startup phase ([4], [5], and references therein). In particular, Macchi and Eweda have proven [5] that the DD algorithm converges to the optimum tap setting in the noiseless case, whenever the initial tap setting corresponds to an open eye pattern. In addition, some capabilities for blind convergence, although weak, are exhibited by the algorithm with a closed-eye condition, leading to certain stationary points for which domains of attraction can be defined [5].

Based on the above considerations, it is worth exploring the possibility of retaining the advantages of simplicity and smoothness of the DD algorithm while attempting to substantially improve its blind convergence capabilities. The basic idea in this paper is that this can be achieved by stopping adaptation when the reliability (in a probabilistic sense) of the self-decided output error is not high enough. More precisely, an easy-to-generate binary-valued flag tells the equalizer whether the output error on the current decision may be reliably used in the standard DD algorithm. If not, adaptation is stopped for that iteration.

The above "stop-and-go" operation mode of the standard DD algorithm is an attractive solution of the problem of blind joint equalization and carrier recovery. As a matter of fact, this algorithm has been successfully implemented and tested in a new commercial 64-QAM digital radio system [6].

The paper is organized as follows. In Section II, we show that convergence is achieved for the DD algorithm, provided that a proper fraction of the self-decided errors having incorrect sign are not used for adaptation. A case study shows that the probability of not using an incorrect sign for the output error in the standard DD algorithm should be made only slightly greater than zero before convergence takes place. In Section III, the proposed "stop-and-go" DD algorithm is described achieving this goal. The algorithm may be promptly extended to joint carrier phase recovery, joint automatic gain control (AGC), decision feedback equalizer (DFE—and decided symbols are incorrect!), and can also be used with a clipped pilot vector. Finally, we show in Section IV how the proposed technique is related to the Benveniste–Goursat

Paper approved by the Editor for Digital Communications of the IEEE Communications Society. Manuscript received April 14, 1986; revised March 23, 1987. This work was supported by GTE Telecomunicazioni S.p.A., Italy.

G. Picchi is with the Istituto di Elettronica e Telecomunicazioni, Università di Pisa, I-56100 Pisa, Italy.

G. Prati is with the Dipartimento di Informatica, Sistemistica e Telematica, Università di Genova, I-16145 Genova, Italy.

IEEE Log Number 8716019.

0090-6778/87/0900-0877$01.00 © 1987 IEEE

heuristic algorithm [3], both being modifications of the Bussgang technique for nonlinear blind deconvolution [8].

II. Elaborating on the Decision-Directed Algorithm

Consider a two-dimensional linear equalizer with N complex taps $\boldsymbol{c} = \boldsymbol{c}_R + j\boldsymbol{c}_I$. Let $\boldsymbol{x}_n$ be the complex input vector and z_n the complex equalizer output sample at the time nT, with T being the symbol period

$$z_n = \boldsymbol{x}_n^T \boldsymbol{c}_n \tag{1}$$

(see Fig. 1). The complex output error at time nT is

$$e_n = z_n - a_n \tag{2}$$

where $\hat{a}_n$ is the transmitted complex symbol. At the same time, the estimated error is

$$\hat{e}_n = z_n - \hat{a}_n \tag{3}$$

where $\hat{a}_n$ is the decided complex symbol.

The DD complex adaptation algorithm is[1]

$$\boldsymbol{c}_{n+1} = \boldsymbol{c}_n - \alpha \hat{e}_n \bar{\boldsymbol{x}}_n \tag{4}$$

where α is the real adaptation step size. Algorithm (4) may be split into real and imaginary parts as follows:

$$\begin{aligned} \boldsymbol{c}_{n+1,R} &= \boldsymbol{c}_{n,R} - \alpha(\hat{e}_{n,R}\boldsymbol{x}_{n,R} + \hat{e}_{n,I}\boldsymbol{x}_{n,I}) \\ \boldsymbol{c}_{n+1,I} &= \boldsymbol{c}_{n,I} + \alpha(\hat{e}_{n,R}\boldsymbol{x}_{n,I} - \hat{e}_{n,I}\boldsymbol{x}_{n,R}) \end{aligned} \tag{5}$$

which can also be rewritten as

$$\begin{aligned} \boldsymbol{c}_{n+1,R} = \boldsymbol{c}_{n,R} - \alpha(&\gamma_{n,R} \operatorname{sgn} \hat{e}_{n,R} |e_{n,R}| \boldsymbol{x}_{n,R} \\ &+ \gamma_{n,I} \operatorname{sgn} \hat{e}_{n,I} |e_{n,I}| \boldsymbol{x}_{n,I}) \\ \boldsymbol{c}_{n+1,I} = \boldsymbol{c}_{n,I} + \alpha(&\gamma_{n,R} \operatorname{sgn} \hat{e}_{n,R} |e_{n,R}| \boldsymbol{x}_{n,I} \\ &- \gamma_{n,I} \operatorname{sgn} \hat{e}_{n,I} |e_{n,I}| \boldsymbol{x}_{n,R}) \end{aligned} \tag{6}$$

where $\gamma_{n,R} = |\hat{e}_{n,R}|/|e_{n,R}|$ and $\gamma_{n,I} = |\hat{e}_{n,I}|/|e_{n,I}|$;

Suppose now that the initial tap setting $\boldsymbol{c}_0 = \boldsymbol{c}_{0,R} + j\boldsymbol{c}_{0,I}$ corresponds to the closed-eye condition. With this situation, a large amount of data will be decided incorrectly, that is, $\hat{a}_n \neq a_n$, and convergence will not take place. What may be reasonably assumed to determine the behavior of the DD algorithm (5) are the probabilities

$$p_n = P\{\operatorname{sgn} \hat{e}_{n,R} = \operatorname{sgn} e_{n,R}\} = P\{\operatorname{sgn} \hat{e}_{n,I} = \operatorname{sgn} e_{n,I}\} \tag{7}$$

$$\begin{aligned} q_n = 1 - p_n &= P\{\operatorname{sgn} \hat{e}_{n,R} \neq \operatorname{sgn} e_{n,R}\} \\ &= P\{\operatorname{sgn} \hat{e}_{n,I} \neq \operatorname{sgn} e_{n,I}\} \end{aligned} \tag{8}$$

of correct and incorrect choice of the sign for $|e_{n,R}|$ and $|e_{n,I}|$. Indeed, the ratios $\gamma_{n,R}$, $\gamma_{n,I}$ in (6) may be regarded as stochastic gain factors in the range of 0 to 1 whose average effect could be accounted for by a proper choice of α.

To prove that this is true, we have modified algorithm (5) as follows:

$$\begin{aligned} \boldsymbol{c}_{n+1,R} &= \boldsymbol{c}_{n,R} - \alpha(f_{n,R}\hat{e}_{n,R}\boldsymbol{x}_{n,R} + f_{n,I}\hat{e}_{n,I}\boldsymbol{x}_{n,I}) \\ \boldsymbol{c}_{n+1,I} &= \boldsymbol{c}_{n,I} + \alpha(f_{n,R}\hat{e}_{n,R}\boldsymbol{x}_{n,I} - f_{n,I}\hat{e}_{n,I}\boldsymbol{x}_{n,R}) \end{aligned} \tag{9}$$

where $f_{n,R}$ and $f_{n,I}$ are independent random variables taking on values equal to 0 or 1 also independent of e_n and $\boldsymbol{x}_n$ for which the following probabilities are defined:

$$\begin{aligned} p_{sc} = P\{\text{stop}|\text{no error}\} &= P\{f_{n,R} = 0 | \operatorname{sgn} \hat{e}_{n,R} \\ &= \operatorname{sgn} e_{n,R}\} = P\{f_{n,I} = 0 | \operatorname{sgn} \hat{e}_{n,I} = \operatorname{sgn} e_{n,I}\} \end{aligned} \tag{10}$$

[1]Overbar denotes complex conjugate.

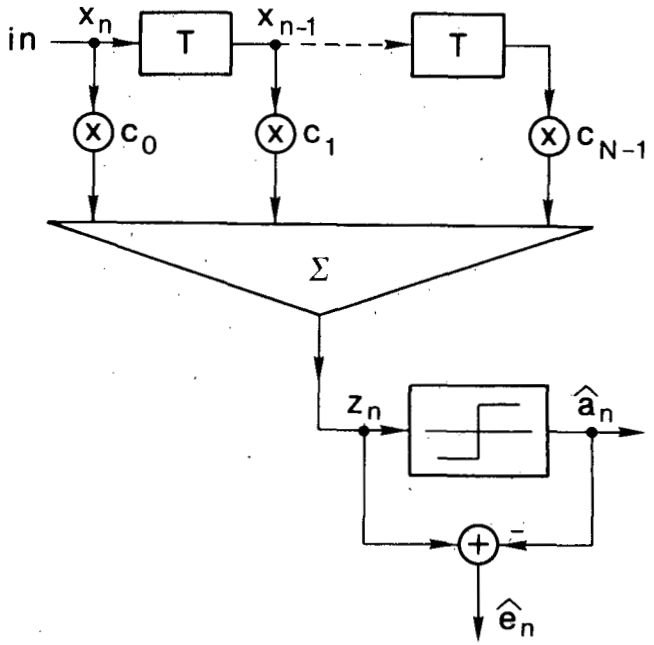


Fig. 1. Basic two-dimensional linear equalizer.

$$\begin{aligned} p_{se} = P\{\text{stop}|\text{error}\} &= P\{f_{n,R} = 0 | \operatorname{sgn} \hat{e}_{n,R} \\ &\neq \operatorname{sgn} e_{n,R}\} = P\{f_{n,I} = 0 | \operatorname{sgn} \hat{e}_{n,I} \neq \operatorname{sgn} e_{n,I}\}. \end{aligned} \tag{11}$$

This case study consists of assuming that the error e_n is available, and is used to stop adaptation ($f_{n,R}$ or $f_{n,I}$ set to 0) with a given probability p_{se} imposed from outside and independent of n when $\hat{e}_{n,R}$ or $\hat{e}_{n,I}$ have incorrect sign.

A typical plot of convergence curves for this case with $p_{sc} = 0$ is shown in Fig. 3 for a five-tap equalizer.[2] The curve for $p_{se} = 0$ refers to the DD algorithm in the closed-eye condition.

Surprising, but not that much, values of p_{se} on the order of 0.1 are enough to make the DD algorithm converge. The larger p_{se}, the higher the speed of convergence, of course. Making $p_{sc} \neq 0$ corresponds to possibly stopping adaptation when it would not be necessary to. It is intuitive, however, that this mainly has an effect on the speed of convergence rather than on the convergence itself, at least for small values of p_{sc}. This effect is shown in Fig. 4 where a sample curve from Fig. 3 for $p_{se} = 0.2$, $p_{sc} = 0$ is reported, and a set of sample curves for $p_{sc} = 0.05$, 0.06, and 0.1, also with $p_{se} = 0.2$, is superimposed. Convergence is maintained, but speed of convergence is reduced.

Based on this evidence, we propose in the next section a blind operation mode for the DD algorithm which achieves considerably high values of $p_{se,n}$ without sacrificing too much speed of convergence ($p_{sc,n}$ small).

III. The "Stop-and-Go" Decision-Directed Algorithm

In the preceding section, we have verified that the DD algorithm (5), modified as in (9), converges provided that either flag $f_{n,R}$ or $f_{n,I}$ is set to zero with a sufficiently high probability p_{se} whenever the sign of $\hat{e}_{n,R}$ or $\hat{e}_{n,I}$ is incorrect. The action taken in Section II toward this end was based on the knowledge of e_n. For the "plain" DD algorithm, the region R of the z_R axis where the conditional event $\{\operatorname{sgn} \hat{e}_{n,R} = \operatorname{sgn} \hat{e}_{n,R} | a_{n,R} = s_i\}$ occurs is shown in Fig. 5(a) for the transmitted symbol $s_i = 3$ in a 64-QAM constellation plane. Similar patterns are found for the other symbols.

Looking at the pattern in Fig. 5(a), it is reasonable to anticipate that an effective strategy for blind convergence

[2]The channel impulse response used here and throughout the paper is shown in Fig. 2, and modulation format is 64-QAM (except for Fig. 8).

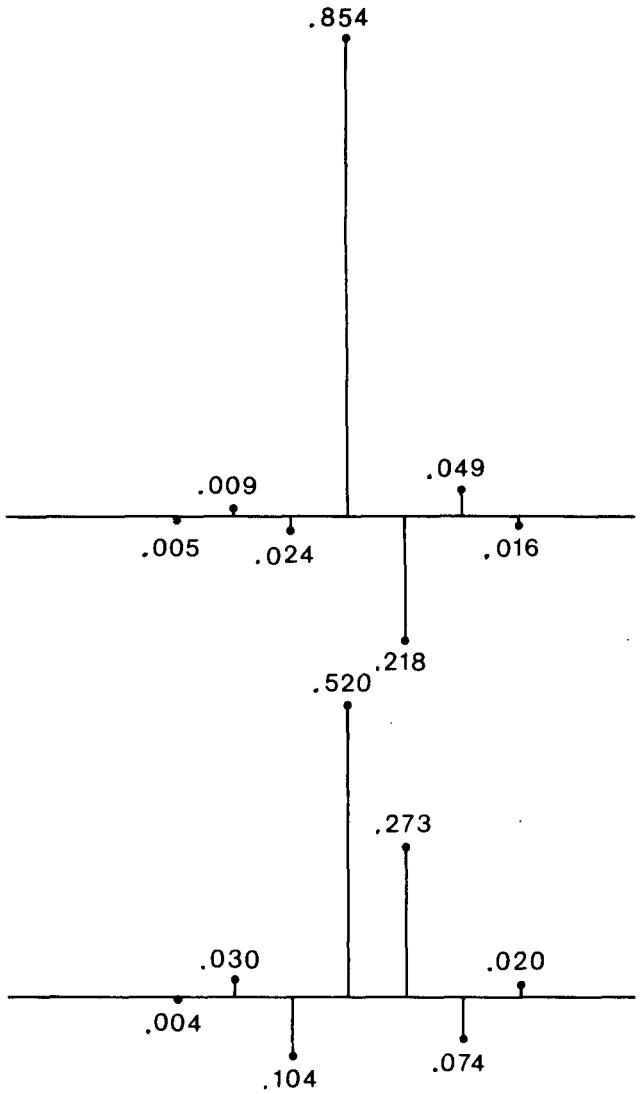


Fig. 2. Channel impulse response.

should be based on retaining certain areas favorable on the average for adaptation and excluding certain others. For instance, the maximum level error (MLE) technique [7] uses for adaptation only those $z_{n,R}$ (or $z_{n,I}$) such that $|z_{n,R}| > 7(|z_{n,I}| > 7)$ in Fig. 5(a). These values of the received signal have the property that the corresponding error with respect to the decided symbol almost surely has the sign of the true error. In all other cases, adaptation is stopped. Due to the low probability of the events $\{|z_{n,R}| > 7\}$ and $\{|z_{n,I}| > 7\}$ (high values of p_{sc}), the speed of convergence of the MLE algorithm is considerably low.

Another technique for blind convergence, already mentioned in the Introduction, was proposed by Sato [1] and studied by Benveniste and Goursat [3]. With this approach, the M-ary alphabet of the transmitted symbols is roughly considered to be binary when the decisions on the M-ary symbols are not reliable in the sense that the errors to be used in the adaptation algorithm are taken with respect to a center-of-gravity point on each axis rather than to the decisions. With this algorithm, convergence is good, but not very smooth both in the transient and in the steady-state period due to the "wildness" of the Sato error when compared to the true error.

In this section, a "stop-and-go" DD algorithm similar to the case study of Section II is presented, which is not based on the knowledge of e_n, of course. In this algorithm, two Sato-like errors

$$\tilde{e}_{n,R} = z_{n,R} - (\operatorname{sgn} z_{n,R})\beta_n$$
$$\tilde{e}_{n,I} = z_{n,I} - (\operatorname{sgn} z_{n,I})\beta_n \tag{12}$$

are generated, with β_n being a suitable real value possibly changing with n, and used to determine on which intervals of the z_R and z_I axes the error on the decided symbol may be used for adaptation. More specifically, the DD algorithm (9) now uses the following flags:

$$f_{n,R} = \begin{cases} 1 & \text{if } \operatorname{sgn} \hat{e}_{n,R} = \operatorname{sgn} \tilde{e}_{n,R} \\ 0 & \text{if } \operatorname{sgn} \hat{e}_{n,R} \neq \operatorname{sgn} \tilde{e}_{n,R} \end{cases}$$
$$f_{n,I} = \begin{cases} 1 & \text{if } \operatorname{sgn} \hat{e}_{n,I} = \operatorname{sgn} \tilde{e}_{n,I} \\ 0 & \text{if } \operatorname{sgn} \hat{e}_{n,I} \neq \operatorname{sgn} \tilde{e}_{n,I}. \end{cases} \tag{13}$$

When the event $[\operatorname{sgn} \hat{e}_{n,R} = \operatorname{sgn} \hat{e}_{n,R}\}$ occurs and the choice of β_n is proper, the conditional probability

$$P_{n|\text{go}} = P\{\text{no error}|\text{go}\} = P\{\operatorname{sgn} \hat{e}_{n,R} = \operatorname{sgn} e_{n,R} | \operatorname{sgn} \hat{e}_{n,R} = \operatorname{sgn} \tilde{e}_{n,R}\} = P\{\operatorname{sgn} \hat{e}_{n,I} = \operatorname{sgn} e_{n,I} | \operatorname{sgn} \hat{e}_{n,I} = \operatorname{sgn} \tilde{e}_{n,I}\} \tag{14}$$

is sufficiently high, and $\hat{e}_{n,R}$ may be confidently used in the DD algorithm. The two flags $f_{n,R}$ and $f_{n,I}$ therefore restrict operation of algorithm (9) to a region of higher reliability. Fig. 5(b) shows the operation intervals (dashed) for each axis when $\beta_n = 6$. The intersection of the dashed regions of Fig. 5(a) and (b) indicates where the algorithm operates correctly (for $a_{n,R} = 3$). It is apparent that the correct operation region depends on the transmitted symbol.

It can be shown by a direct measure of some performance parameters, discussed below, that a proper choice of β_n can be made, favorable *on the average*. Therefore, β_n is used as a reference point to identify regions on each axis where $\hat{e}_{n,R}$ and $\hat{e}_{n,I}$ may be used more confidently. The conditional probability $p_{n|\text{go}}$ and its complement

$$q_{n|\text{go}} = P\{\text{error}|\text{go}\} = 1 - p_{n|\text{go}} \tag{15}$$

correspond in a sense to p_n and q_n in (7) and (8) for the present case. However, performance is better related to the probabilities (real axis only, for simplicity)

$$p_{n,\text{go}} = P\{\text{no error, go}\} = P\{\operatorname{sgn} \hat{e}_{n,R} = \operatorname{sgn} e_{n,R}, f_{n,R} = 1\} = P\{\operatorname{sgn} \hat{e}_{n,R} = \operatorname{sgn} e_{n,R}, \cdot \operatorname{sgn} \hat{e}_{n,R} = \operatorname{sgn} \tilde{e}_{n,R}\} = (1 - p_{sc})p_n$$
$$q_{n,\text{go}} = P\{\text{error, go}\} = P\{\operatorname{sgn} \hat{e}_{n,R} \neq \operatorname{sgn} e_{n,R}, f_{n,R} = 1\} = P\{\operatorname{sgn} \hat{e}_{n,R} \neq \operatorname{sgn} e_{n,R}, \cdot \operatorname{sgn} \hat{e}_{n,R} = \operatorname{sgn} \tilde{e}_{n,R}\} = (1 - p_{se})q_n. \tag{17}$$

In fact, $g_{n|\text{go}}$ may be very small, as desired for high reliability, but $p_{n,\text{go}}$ may also turn out to be very small, leading to very slow convergence as in the MLE algorithm. It is therefore the combination of $q_{n|\text{go}}$, $p_{n,\text{go}}$, and $g_{n,\text{go}}$ that determines the behavior of the algorithm.

Notice that p_n and q_n in (7) and (8) for the DD algorithm (from now on referred to as $p_{n,\text{DD}}$ and $q_{n,\text{DD}}$) and p_n and q_n in (16) and (17) for the present case may have different values, as $p_{se} \neq 0$ and $p_{sc} \neq 0$ affect the statistics of the output error $\hat{e}_n$.

A proper choice of β_n causes $q_{n,\text{go}}$ to be sufficiently smaller than $q_{n,\text{DD}}$, that is, adaptation stops with probability p_{se} when a sign error occurs for a component of $\hat{e}_n$. This makes the DD algorithm converge. When using flags (13), which are not based on e_n, the probability $q_{n,\text{go}}$ cannot be made smaller than

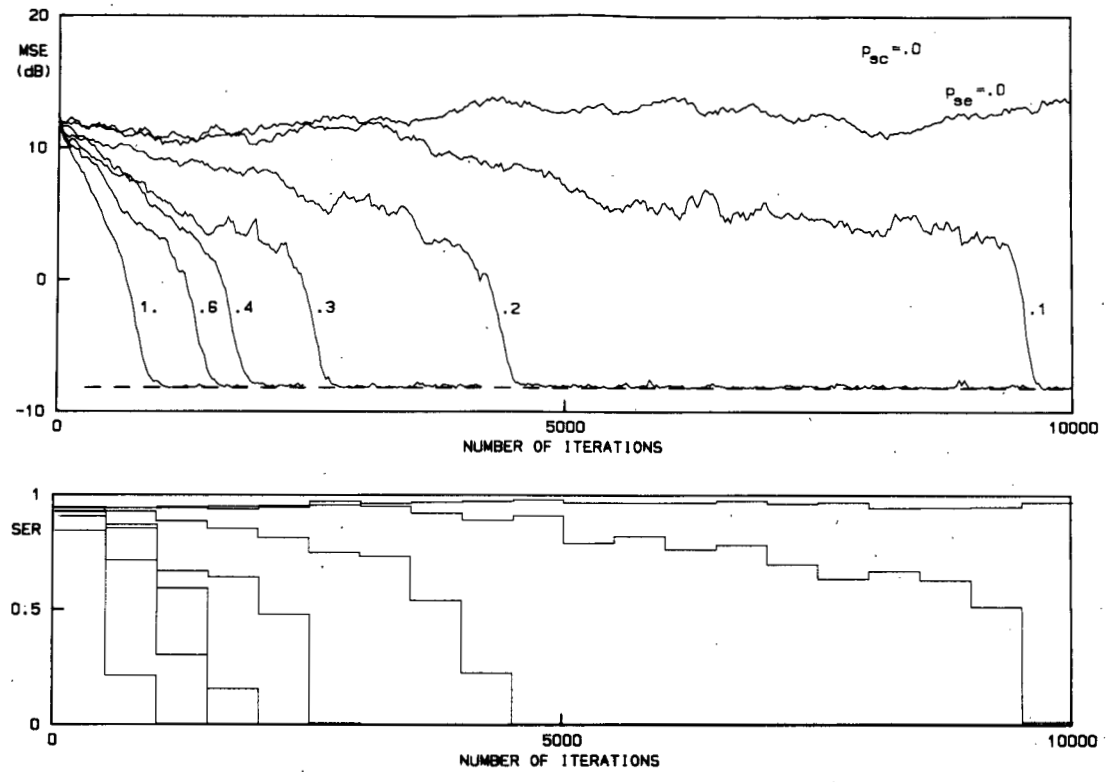


Fig. 3. MSE (top) and measured symbol error rate (bottom) versus number of iterations for the case study of Section II for various values of p_{se} with $p_{sc} = 0$. Other parameters are $\alpha = 5 \times 10^{-4}$, SNR = 60 dB.

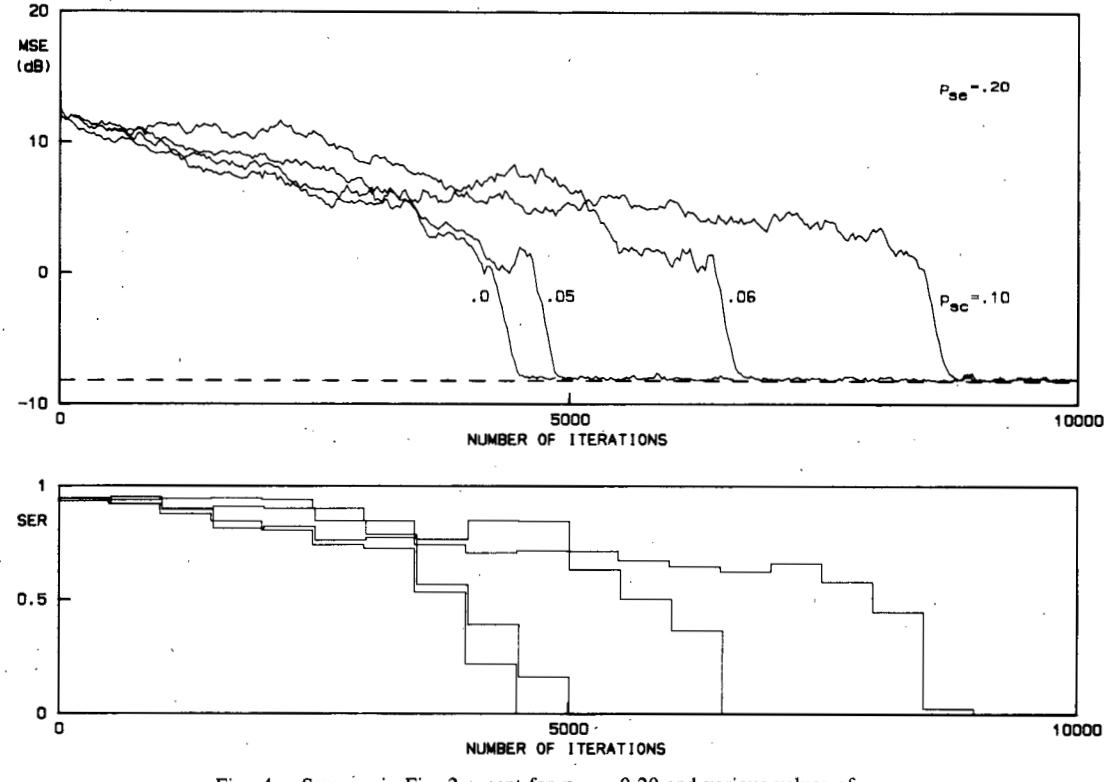


Fig. 4. Same as in Fig. 2 except for $p_{se} = 0.20$ and various values of p_{sc}.

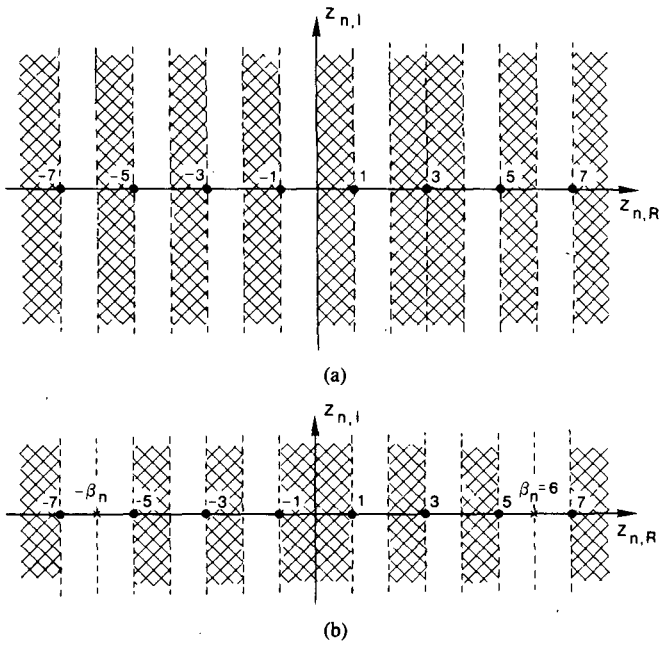


Fig. 5. Regions of the 64-QAM constellation plane where (a) the event $\{\text{sgn } \hat{e}_{n,R} = \text{sgn } \hat{e}_{n,R} | a_{n,R} = 3\}$ occurs; (b) the event $\{\text{sgn } \hat{e}_{n,R} = \text{sgn } \tilde{e}_{n,R}\}$ occurs when $\beta_n = 6$.

$q_{n,\text{DD}}$ without causing $p_{n,\text{go}} < p_{n,\text{DD}}$ during convergence. In other words, the probability p_{sc} in (10) that during convergence the flags $f_{n,R}$ or $f_{n,I}$ be set to zero when no sign error occurs cannot be made equal to zero. This affects the speed of convergence as discussed in Section II.

Fig. 6(a)–(d) show measured values of $p_{n,\text{go}}$, $q_{n,\text{go}}$, and $q_{n|\text{go}}$ versus n for the proposed "stop-and-go" DD algorithm (9) with flags as in (13) for various values of $\beta_n = \beta$ in the range of 4–8. Also shown are

$$P\{\text{go}\} = p_{n,\text{go}} + q_{n,\text{go}} \tag{18}$$

and

$$p_{se} = P\{\text{stop}|\text{error}\} = P\{\text{sgn } \hat{e}_{n,R} \neq \text{sgn } \tilde{e}_{n,R} | \text{sgn } \hat{e}_{n,R} \neq \text{sgn } e_{n,R}\}. \tag{19}$$

When convergence occurs, $P\{\text{sgn } \hat{e}_{n,R} \neq \text{sgn } e_{n,R}\}$ equals zero and the probability in (19) is conventionally set to 1 in Fig. 6(a)–(d). An effective algorithm should exhibit fast convergence of $q_{n,\text{go}}$ and $q_{n|\text{go}}$ to zero and $p_{n,\text{go}}$ as large as possible. It is apparent that the choice of β_n is crucial in this regard. In principle, an optimum value of β_n exists for each n. In a practical implementation, however, the choice of a fixed $\beta_n = \beta$ is more convenient. It was found by simulation that the value $\beta = 6$ yields the best compromise for 64-QAM between speed of convergence and smooth steady state for all channels of practical interest examined. Fig. 7(a) shows a set of ten sample convergence curves for the reference channel of Fig. 2 when using algorithm (9) with flags as in (13) and $\beta = 6$. No phase offset was assumed in this example. For a comparison, performance of the Sato/Benveniste–Goursat algorithm is shown in Fig. 7(b). Numerical results refer to the 64-QAM format, as said. To test convergence for longer equalizers, a 32-tap "stop-and-go" blind equalizer has also been simulated for the telephone channel given in [3] with a 16-QAM format. Simulation results are reported in Fig. 8(a). Fig. 8(b) refers again to the Sato/Benveniste–Goursat algorithm.

Extending to Carrier Phase Recovery and Automatic Gain Control (AGC)

The above "stop-and-go" technique may be readily extended to carrier phase recovery. With $f_{n,R}$ and $F_{n,I}$ as in (13),

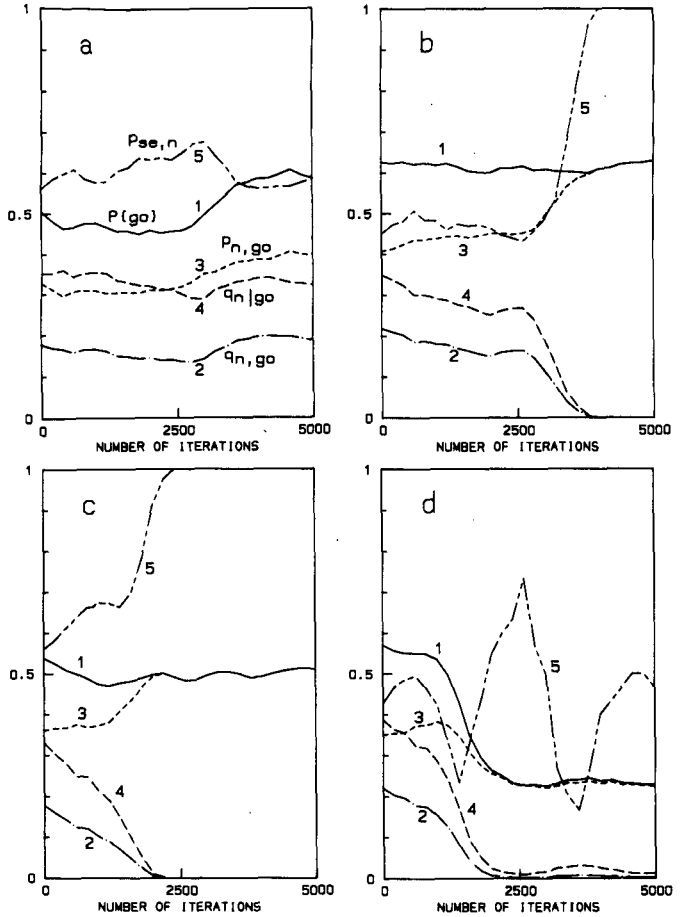


Fig. 6. Measured values of $P\{\text{go}\}$ (1), $q_{n,\text{go}}$ (2), $p_{n\text{go}}$ (3), $q_{n|\text{go}}$ (4), and $p_{se,n}$ (5) for (a) $\beta = 4$, (b) $\beta = 5$, (c) $\beta = 6$, and (d) $\beta = 8$ for the "stop-and-go" algorithm.

the classic decision-directed MSE iterative algorithm for carrier phase recovery [7] becomes

$$\Phi_{n+1} = \Phi_n - \mu(-f_{n,R}\hat{e}_{n,R}z_{n,I} + f_{n,I}\hat{e}_{n,I}z_{n,R}). \tag{20}$$

In practical digital radio systems, an analog AGC is often inserted before carrier phase recovery and the baseband equalizer to compensate for possible amplitude variations of the incoming signal. In this case, the reference tap of the equalizer is normally set to 1. The cascade of AGC, carrier phase recovery, and equalizer is shown in Fig. 9. Extension of the "stop-and-go" DD MSE adaptation algorithm to this "external" real gain G_n leads to the following recursion:

$$G_{n+1} = G_n - \mu_A(f_{n,R}\hat{e}_{n,R}z_{n,R} + f_{n,I}\hat{e}_{n,I}z_{n,I}). \tag{21}$$

Sample curves for the blind convergence of the cascade of AGC, carrier phase recovery, and linear equalizer, with center tap initially set to 1, are shown in Fig. 10. Fig. 11 refers to the case of convergence in the presence of a carrier frequency offset $\Delta f/R = 10^{-4}$, with R being the baud rate. In the present example, this is the maximum Δf for which convergence has been obtained.

Extending to Decision Feedback Equalizer (DFE)

Extension of any blind equalization technique to a decision feedback scheme appears to be difficult to justify in terms of cancellation of interferers, as is usually done when decisions are correct. A logic procedure to derive a "stop-and-go" blind DFE consists of first considering the corresponding linear recursive version and then replacing the input samples of the decision device with the corresponding output symbols. The

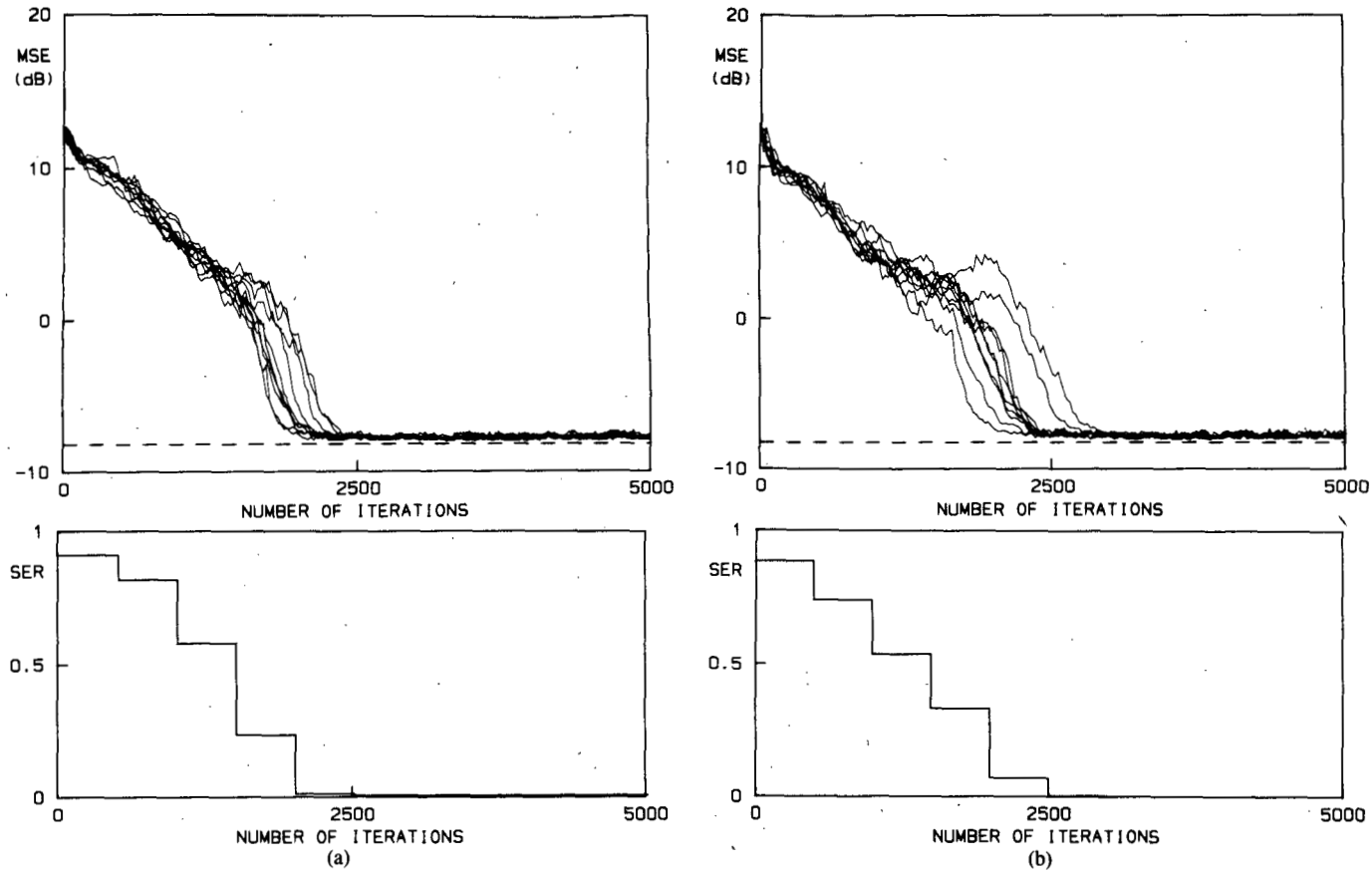


Fig. 7. Convergence of MSE (top) and measured symbol error rate (bottom) versus number of transmitted symbols for a five-tap "stop-and-go" linear equalization algorithm. Parameters are $\beta = 6$, $\alpha = 6 \times 10^{-4}$, SNR = 60 dB.(b) Same as in (a) for the Sato/Benveniste-Goursat equalization algorithm. Parameters are $\beta = 5.25$, $k_1 = 4$, $k_2 = 2$, $\alpha = 5 \times 10^{-5}$, SNR = 60 dB.

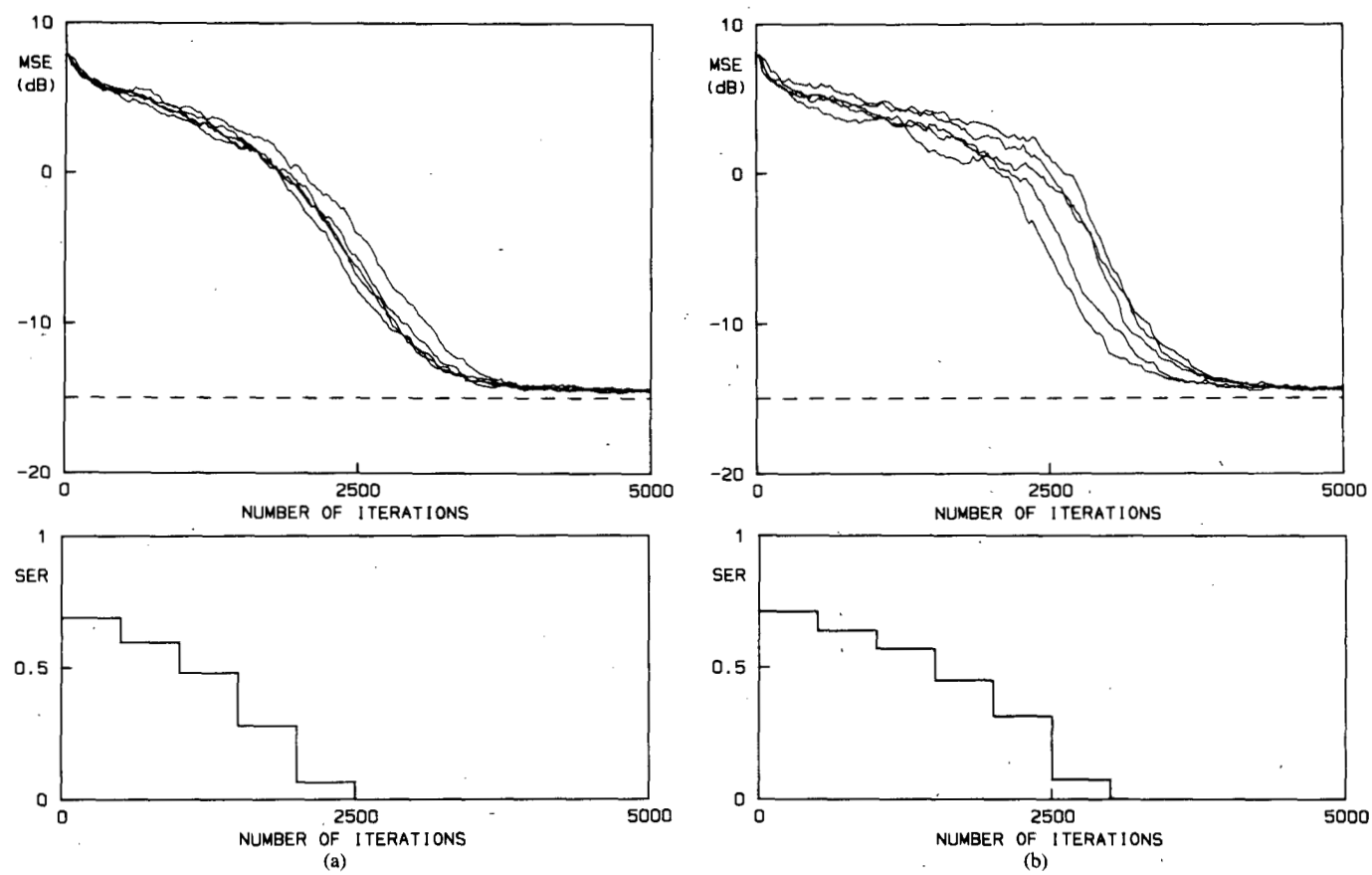


Fig. 8. (a) Same as in Fig. 7(a) for a 32-tap equalizer for the telephone channel given in [3] with 16-QAM format. Parameters are $\beta = 6$, $\alpha = 4 \times 10^{-4}$, SNR = 27 dB. (b) Same as in (a) for the Sato/Benveniste-Goursat algorithm. Parameters are $\beta = 5.25$, $k_1 = 4$, $k_2 = 2$, $\alpha = 5 \times 10^{-5}$, SNR = 27 dB.

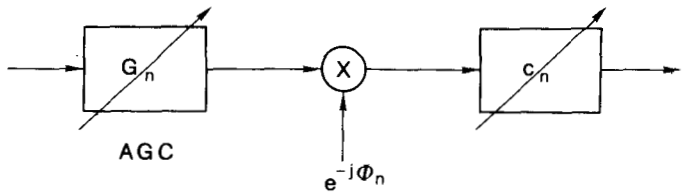


Fig. 9. Block diagram of the cascade of automatic gain control (AGC), carrier phase recovery, and adaptive equalizer.

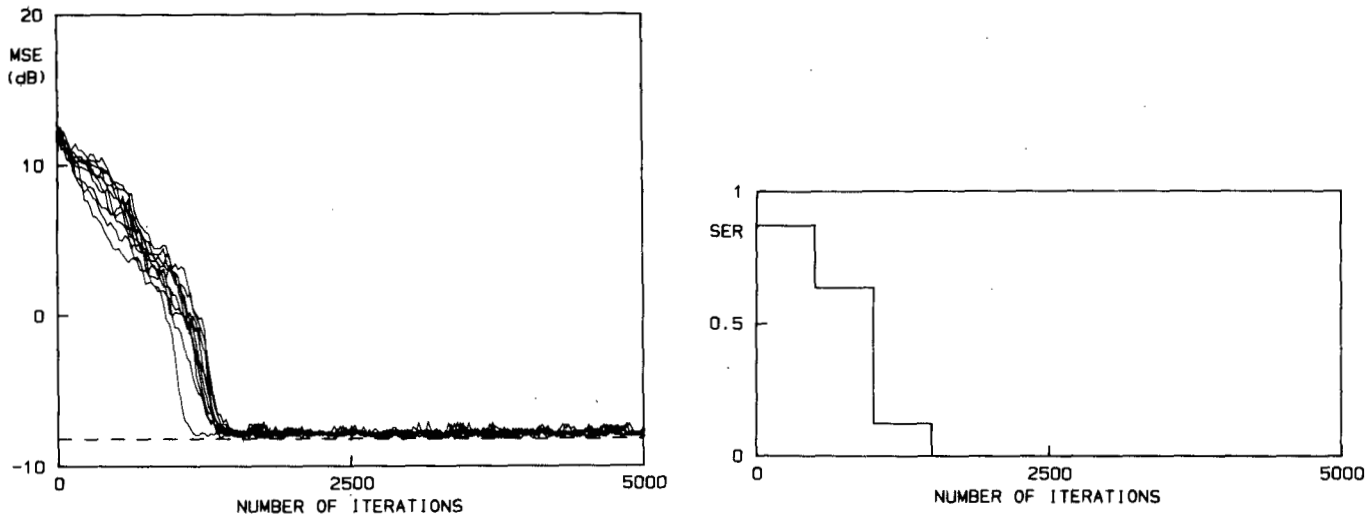


Fig. 10. Same as in Fig. 7(a) for the joint blind adaptation of AGC, carrier recovery, and linear equalizer. Parameters are $\beta = 6$, $\alpha = 6 \times 10^{-4}$, $\mu = 1.2 \times 10^{-3}$, $\mu_A = 1.2 \times 10^{-3}$, SNR = 60 dB.

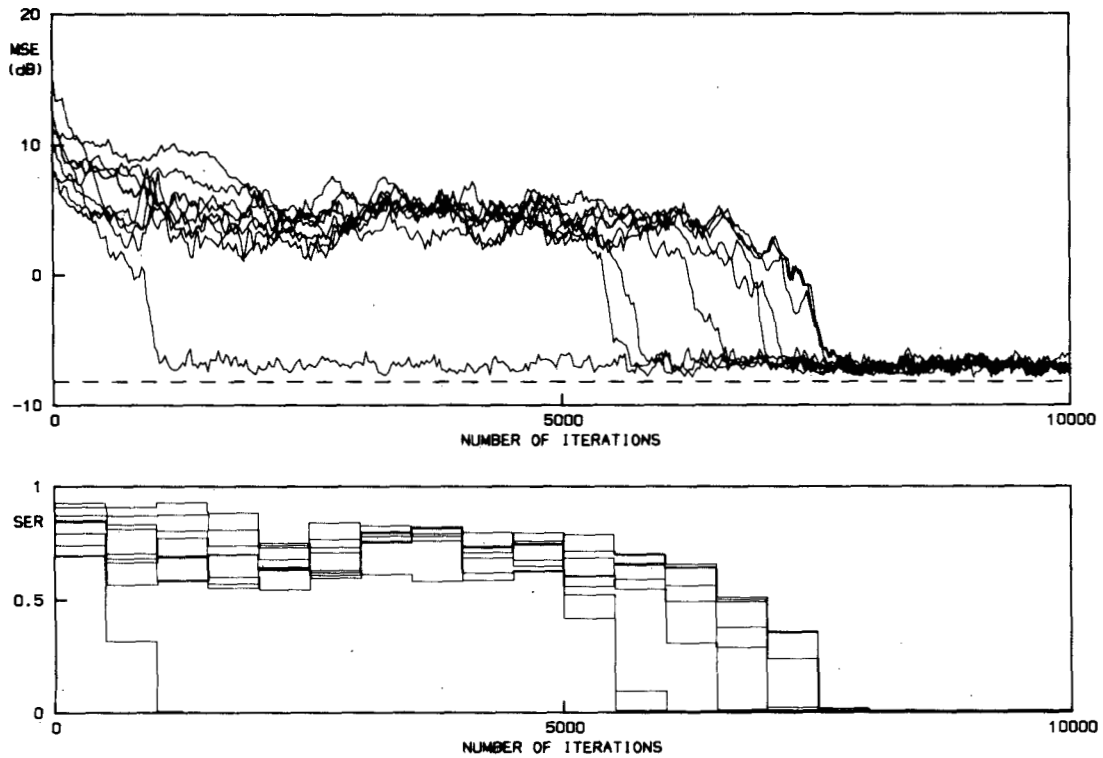


Fig. 11. Same as in Fig. 9 in the presence of a frequency offset $\Delta f/R = 10^{-4}$. R is the baud rate. Parameters are as in Fig. 9.

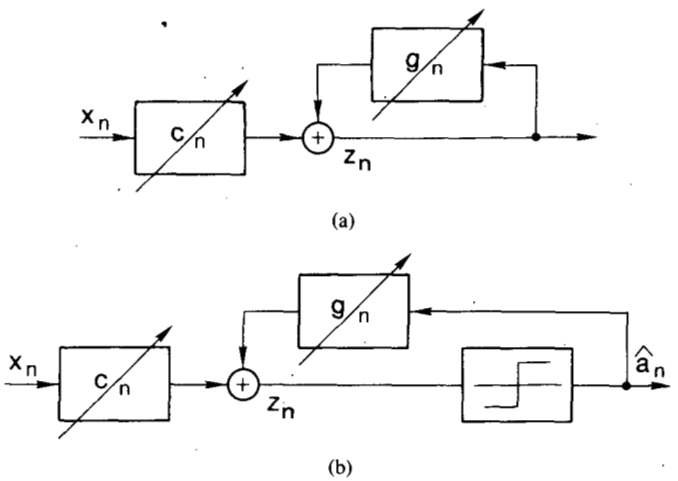


Fig. 12. Block diagram of (a) the two-dimensional linear recursive equalizer, (b) the two-dimensional decision-feedback equalizer.

decided symbols are therefore to be interpreted as approximations of the quantities required in the linear recursive blind equalizer. With this approach, let $\boldsymbol{g}_n$ be the tap coefficients vector for the feedback section of the blind DFE in Fig. 12. By now interpreting $\hat{a}_{n-1}$ as an approximated version of z_{n-1}, independently of whether $\hat{a}_{n-1} = a_{n-1}$ or $\hat{a}_{n-1} \neq a_{n-1}$, we obtain

$$\begin{aligned} \boldsymbol{g}_{n+1,R} &= \boldsymbol{g}_{n,R} - \delta(f_{n,R}\hat{e}_{n,R}\hat{\boldsymbol{a}}_{n-1,R} + f_{n,I}\hat{e}_{n,I}\hat{\boldsymbol{a}}_{n-1,I}) \\ \boldsymbol{g}_{n+1,I} &= \boldsymbol{g}_{n,I} + \delta(f_{n,R}\hat{e}_{n,R}\hat{\boldsymbol{a}}_{n-1,I} - f_{n,I}\hat{e}_{n,I}\hat{\boldsymbol{a}}_{n-1,R}) \end{aligned} \tag{22}$$

with

$$\hat{\boldsymbol{a}}_{n-1} = [\hat{a}_{n-1}\hat{a}_{n-2} \cdots]^T \tag{23}$$

and flags as in (12). The adaptation algorithm for the forward section is again given by (9).

In the above blind DFE, the decided symbols are thus fed back independently of the symbol error rate. Performance of the proposed blind DFE with AGC and carrier recovery is shown in Fig. 13 for $\beta = 6$ and for the same overall number of taps as for the linear nonrecursive equalizer of Fig. 10. Convergence is fully satisfactory, with a lower residual MSE, as expected.

Using a Clipped Pilot Vector

We have seen that the "stop-and-go" technique for the DD algorithm may be applied to joint carrier recovery, AGC, and DFE. Blind convergence was achieved by properly selecting the regions of operation on the (z_R, z_I) plane when the decided error $\hat{e}_n$ is to be used in conjunction with the "pilot vector" $\boldsymbol{v}_n = \boldsymbol{x}_n$. Therefore, all results in the preceding sections are to be intended "for a given probability density of the pilot vector $\boldsymbol{v}_n$."

Suppose now that the pilot vector

$$\boldsymbol{v}_n = [v_{n,0}v_{n,1} \cdots v_{n,N-1}]^T \tag{24}$$

is a nonlinear function of $\boldsymbol{x}_n$ such as

$$v_{n,j} = \operatorname{sgn} x_{n,j} \tag{25}$$

or

$$v_{n,j} = \begin{cases} \operatorname{sgn} x_{n,j} & \text{if } |x_{n,j}| > 1 \\ x_{n,j} & \text{if } |x_{n,j}| < 1 \end{cases} \tag{26}$$

(clipped pilot vector). Direct use of a clipped pilot vector in recursions (5) leads to DD adaptation algorithms which exhibit surprisingly strong blind convergence capabilities. Fig. 14 shows, for the DD algorithm and for the reference channel used throughout the paper, ten sample convergence curves versus number of iterations for $\boldsymbol{v}_n = \operatorname{sgn} \boldsymbol{x}_n$ (hard limiter). The DD algorithm converges, and the speed of convergence is remarkable. It can be verified that $p_{n,\mathrm{DD}}$ in (6), not large enough when e_n is used with $\boldsymbol{x}_n$, is sufficiently large when the statistics of the components of $\boldsymbol{x}_n$ are changed by clipping as in (25).

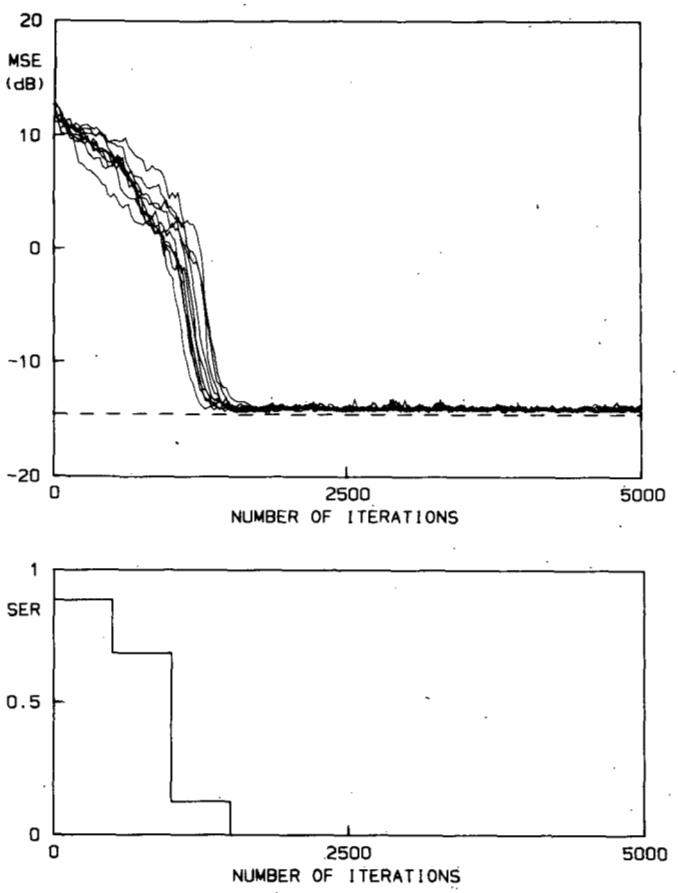


Fig. 13. Same as in Fig. 7(a) for the joint blind adaptation of AGC, carrier recovery, and (3 + 2) − tap DFE. Parameters are $\beta = 5.25$, $\alpha = \delta = 6 \times 10^{-4}$, $\mu = \mu_A = 1.2 \times 10^{-3}$, SNR = 60 dB.

Clipping the pilot vector may be effectively combined with the "stop-and-go" technique by replacing in (9) $\boldsymbol{x}_n$ by $\boldsymbol{v}_n$.

Performance of algorithm (9) using the hard-limiter (25) is shown in Fig. 15. This algorithm is very attractive, combining the simplicity of the pilot vector with the effectiveness of the "stop-and-go" blind algorithm. Best values of β are found in the range of 4–6, depending upon the channel impulse response.

When the soft limiter is used, that is, $\boldsymbol{v}_n$ is as in (26), convergence of algorithm (9) is even faster. Sample curves for this case are shown in Fig. 16 where convergence is achieved in about 500 symbols, as opposed to about 1200 symbols for the nonclipped pilot vector case. The reference error was taken with respect to $\beta = 6$. The residual MSE for this value of β is somewhat higher than in the previous examples. If speed of convergence is of primary concern, a smaller step size α should be used in the steady state.

IV. Relating the "Stop-and-Go Algorithm" to the Bussgang Technique

The Bussgang technique is an iterative blind deconvolution technique applicable to the identification of a channel impulse response, which uses the known (non-Gaussian) probability density of the zero-mean independent identically distributed

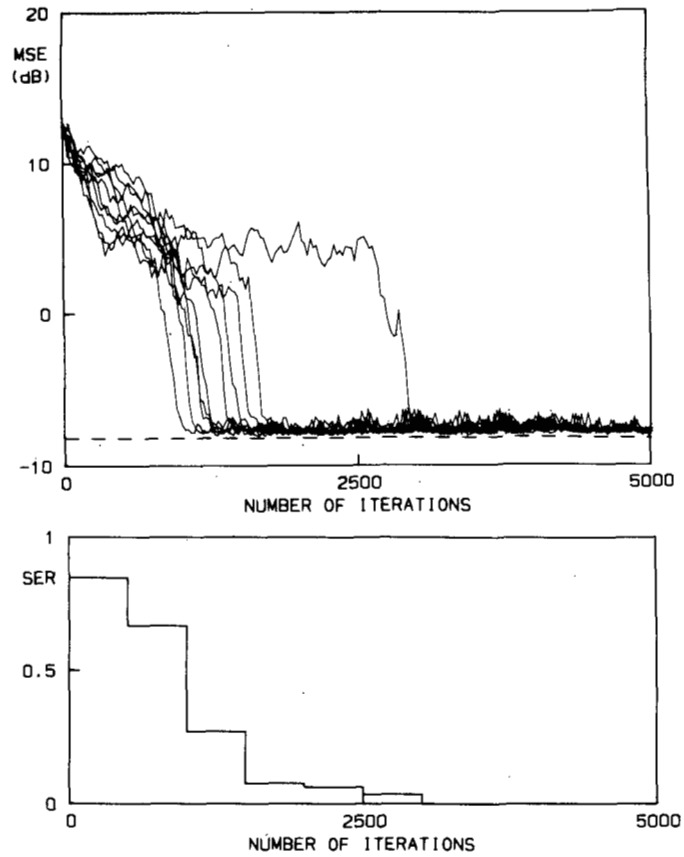


Fig. 14. Convergence of MSE (top) and measured symbol error rate (bottom) versus number of transmitted symbols for the simple estimated-gradient decision-directed algorithm with clipped pilot vector (hard limiter), five-tap linear equalizer with AGC and phase recovery. Parameters are $\alpha = 3 \times 10^{-3}$, $\mu = \mu_A = 6 \times 10^{-3}$, SNR = 60 dB.

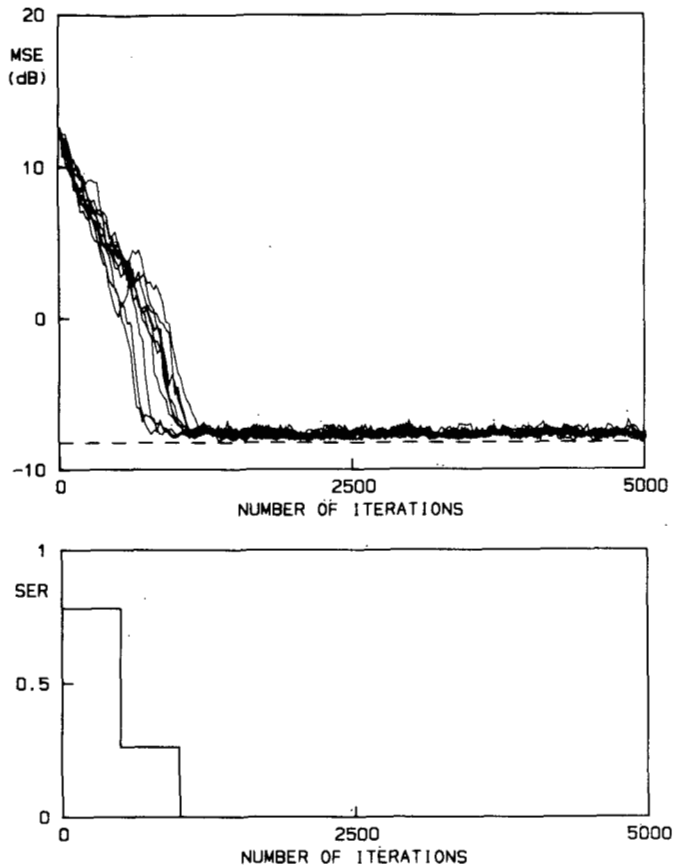


Fig. 15. Same as in Fig. 14 for the "stop-and-go" algorithm. Parameters are $\beta = 4$, $\alpha = 3 \times 10^{-3}$, $\mu = \mu_A = 6 \times 10^{-3}$, SNR = 60 dB.

$\{a_n\}$ and the (approximately) deconvolved output sequence $\{z_n\}$ [8].

In this section, we review briefly the Bussgang technique, showing how it can provide a conceptual framework for the "stop-and-go" algorithm as well as for the MLE, Sato, and Benveniste–Goursat algorithms.

Therefore, let $\boldsymbol{c}_{\text{opt}}^{(j)}$ be the tap vector approximating the inverse filter at the jth Bussgang iteration. The corresponding output sequence $\{z_n^{(j)}\}$ may be written from (2)

$$z_n^{(j)} = \boldsymbol{x}_n^T \boldsymbol{c}_{\text{opt}}^{(j)} = a_n + e_n^{(j)} \tag{27}$$

with $e_n^{(j)}$ now being regarded as the "deconvolution noise" introduced by $\boldsymbol{c}_{\text{opt}}^{(j)}$, which deconvolves the channel only approximately. The sequence $\{e_n^{(j)}\}$ is modeled as Gaussian and independent of the non-Gaussian input data sequence $\{a_n\}$. With these assumptions, it can be found that the optimum (in the MSE sense) memoryless estimator of a_n, given $a_n^{(j)}$, is nonlinear:

$$\hat{a}_n^{(j)} = g[z_n^{(j)}] = E\{a_n | z_n^{(j)}\}. \tag{28}$$

For example, in the case of uniform density of one-dimensional $\{a_n\}$, the nonlinearity $g(z)$ in (28) approximates a soft limiter, and the higher the signal-to-interference ratio, the closer the approximation.

Based on the estimates $\{g[z_n^{(j)}]\}$ in (28), the Bussgang technique determines $\boldsymbol{c}_{\text{opt}}^{(j)}$ as the least square solution of the overdetermined system of linear equations:

$$\boldsymbol{x}_n^T \boldsymbol{c}_{\text{opt}}^{(j+1)} = g[z_n^{(j)}] \qquad \forall n. \tag{29}$$

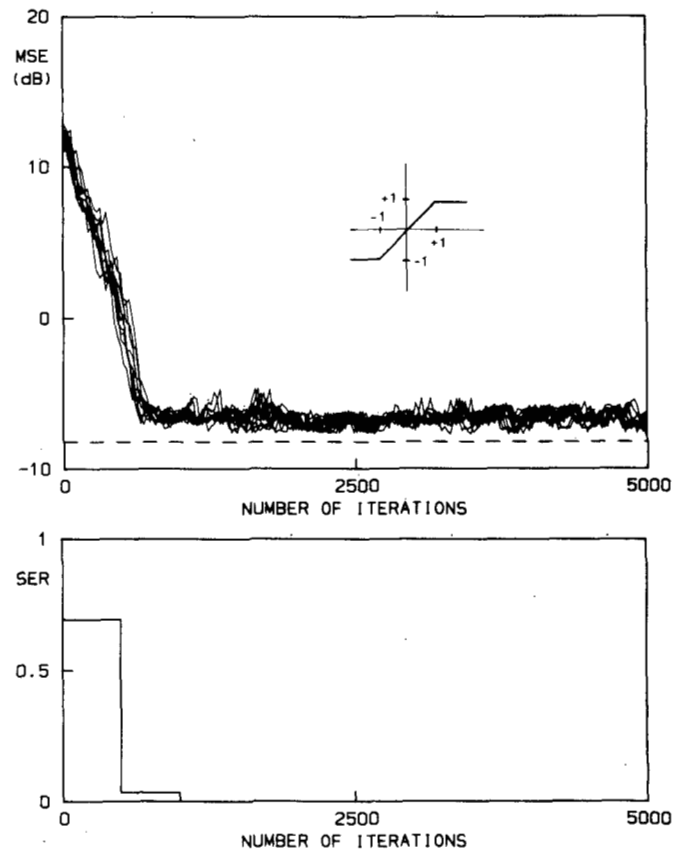


Fig. 16. Same as in Fig. 15 for soft-limited pilot vector. Parameters are $\beta = 6$, $\alpha = 3 \times 10^{-3}$, $\mu = \mu_A = 6 \times 10^{-3}$, SNR = 60 dB.

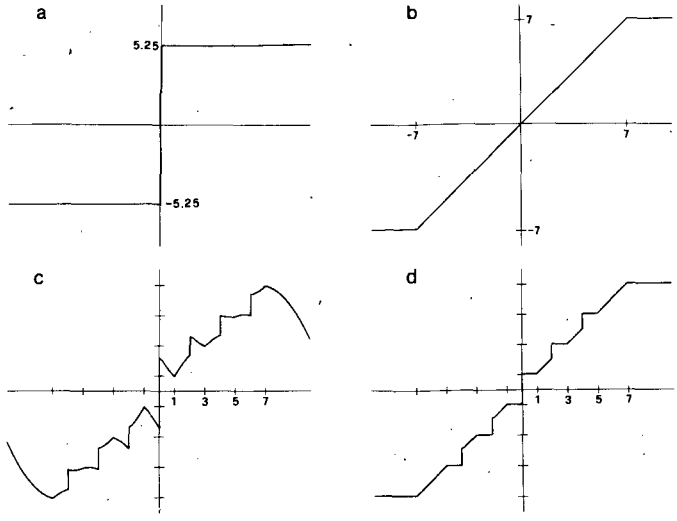


Fig. 17. Nonlinearity $g(t)$ in the one-dimensional 8-ASK case for (a) Sato algorithm, (b) MLE algorithm, (c) Benveniste-Goursat algorithm, and (d) "stop-and-go" DD algorithm.

The least square solution of (29) is the solution of the following equation [9]:

$$\sum_n \bar{x}_n x_n^T c_{\text{opt}}^{(j+1)} = \sum_n g[z_n^{(j)}]\bar{x}_n. \tag{30}$$

The procedure is then repeated, determining $\{\hat{a}_n^{(j+1)}\}$ from $\{z_n^{(j+1)}\}$ in (28) and $c_{\text{opt}}^{(j+2)}$ from (30).

A discussion of the convergence properties of the Bussgang technique is out of the scope of this paper. We refer the interested reader to [8] and references therein. We simply observe here that the iterative Bussgang technique produces a sequence of vectors $\{c_{\text{opt}}^{(j)}\}$ obtained by solving (30), and that the nonlinearity $g(z)$ changes with j, since the statistics of $e_n^{(j)}$ in (27) change with j, until convergence is completed.

The way of solving (30) is unessential to the Bussgang technique. When an iterative method of solution is used for (30), a second iteration loop is nested into the first:

$$c_{k+1}^{(j+1)} = c_k^{(j+1)} - \alpha\bar{x}_k\{x_k^T c_k^{(j+1)} - g[x_k^T c_{\text{opt}}^{(j)}]\} \tag{31}$$

with $c_\infty^{(j+1)} = c_{\text{opt}}^{(j+1)}$. Here, j is the index of the Bussgang outer loop and n is the index for the iterative solution of (30) for a given j. In a practical implementation, the inner loop iterations are necessarily stopped at some $\bar{k} < \infty$, so that

$$c_{\text{opt}}^{(j)} = c_{\bar{k}}^{(j)} \tag{32}$$

$$c_0^{(j+1)} = c_{\bar{k}}^{(j)} \tag{33}$$

which is also a reasonable choice for the initialization of the $(j+1)$th Bussgang iteration. Assuming then that $\bar{k}$ consecutive blocks of incoming samples are used for the Bussgang iterations, recursion (31) be written as follows:

$$c_{k+1}^{(j+1)} = c_k^{(j+1)} - \alpha\bar{x}_{j+k}\{x_{j+k}^T c_k^{(j+1)} - g[x_{j+k}^T c_{\bar{k}}^{(j)}]\},$$
$$k = 0, 1, \cdots, \bar{k}-1. \tag{34}$$

All blind equalization algorithms mentioned in this paper use $\bar{k} = 1$ in (34), implying $k = 0$ for any j, and therefore $j = n$. Using (32) and (33) in (34), we may write

$$c_1^{(n+1)} = c_1^{(n)} - \alpha\bar{x}_n\{x_n^T c_1^{(n)} - g[x_n^T c_1^{(n)}]\} \tag{35}$$

or simply

$$c_{n+1} = c_n - \alpha\check{e}_n\bar{x}_n \tag{36}$$

with

$$\check{e}_n = z_n - g(z_n). \tag{37}$$

Algorithm (36) is therefore a modified Bussgang technique, which appears as an estimated gradient algorithm with the error $\check{e}_n$ defined by (37). Depending on the statistics of $\{a_n\}$ and the choice of $g(z)$, c_n can be made to converge to the solution of

$$E\{\bar{x}_n x_n^T\}c = E\{a_n\bar{x}_n\} \tag{38}$$

as required by the MSE criterion, and the speed of convergence may be remarkably fast.

By comparing (9) to (36), we may easily conclude that the "stop-and-go" algorithm implicitly uses the following nonlinearity:

$$g(z_n) = z_n - \frac{1}{2}[(f_{n,R}+f_{n,I})\hat{e}_n + (f_{n,R}-f_{n,I})\bar{\hat{e}}_n]. \tag{39}$$

The Benveniste-Goursat algorithm uses the following $g(z)$:

$$g(z_n) = z_n - [k_1\hat{e}_n + k_2\tilde{e}_n|\hat{e}_n|]. \tag{40}$$

From (39) and (40), we notice that there are certain similarities between the Benveniste-Goursat algorithm for $k_1 = k_2$ and the "stop-and-go" algorithm. The difference is, however, substantial both conceptually and practically, although performance is similar. From a conceptual viewpoint, the Benveniste-Goursat algorithm is a Sato algorithm heristically modified to reduce undesired effects of a direct use of the Sato error. On the contrary, the "stop-and-go" algorithm is a decision-directed algorithm modified in a very simple manner to improve the reliability of the self-decided output errors used in the adaptation process. From the implementation point of view, the Benveniste-Goursat algorithm is considerably more involved, as it requires performing multibit multiplications as well as modulus operations.

Fig. 17(a)-(d) show $g(z)$ for the Sato, MLE, Benveniste-Goursat, and "stop-and-go" algorithm in the case of one-dimensional 8-ASK modulation format.

V. Conclusion

A decision-directed adaptation algorithm for joint blind convergence of equalizer and carrier phase recovery system

has been presented, which is based on a "stop-and-go" operation mode for the standard DD algorithm. At each iteration, an easy-to-generate binary-valued flag inhibits adaptation if the reliability of the current self-decided output error is not large enough to warrant its use in the adaptation algorithm. This "stop-and-go" technique exalts some blind convergence capabilities that are inherent in the decision-directed algorithm, and may be easily extended to automatic gain control and decision feedback equalization.

REFERENCES

[1] Y. Sato, "A method of self-recovering equalization for multilevel amplitude-modulation systems," *IEEE Trans. Commun.*, vol. COM-23, pp.679-682, June 1975.

[2] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. COM-28, pp. 1867-1875, Nov. 1980.

[3] A. Benveniste and M. Goursat, "Blind equalizers," *IEEE Trans. Commun.*, vol. COM-32, pp. 871-883, Aug. 1984.

[4] J. E. Mazo, "Analysis of decision-directed equalizer convergence," *Bell Syst. Tech. J.*, vol. 59, pp. 1857-1876, Dec. 1980.

[5] O. Macchi and E. Eweda, "Convergence analysis of self-adaptive equalizers," *IEEE Trans. Inform. Theory,* vol. IT-30, pp. 161-176, Mar. 1984.

[6] B. Baccetti *et al.*, "Full digital adaptive equalization in 64QAM radio system," in *Proc. IEEE Int. Conf. Commun. (ICC'86),* Toronto, Canada, June 1986.

[7] R. Yatsuboshi, N. Sata, and K. Aoki, "A convergence of automatic equalizer by maximum level error control," in *Nat. Conv. Rec., IECE Japan,* no. 2192, 1974.

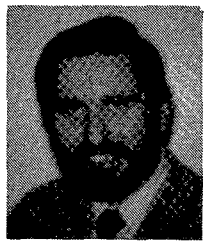
[8] S. Bellini and F. Rocca, "Blind deconvolution: Polyspectra or Bussgang techniques?," in *Digital Communications,* E. Biglieri and G. Prati Eds. North-Holland, 1986, pp. 251-262.

[9] G. Strang, *Linear Algebra and Its Applications.* New York: Academic, 1976, p. 107.

Giorgio Picchi was born in Sarzana, Italy, in 1947. He received the Dr.Ing. degree in electronics engineering from the University of Pisa, Italy, in 1974.

He is Associate Professor in the Institute of Electronics and Telecommunications, University of Pisa, where he is involved in studies on digital signal processing in communications and optical communications systems. Since 1984 he has also been a Research Associate in the Centro di Studio per Metodi e Dispositivi di Radiotrasmissione, CNR, Pisa.

★



Giancarlo Prati (M'80) was born in Rome, Italy, on November 13, 1946. He received the Dr.Ing. degree in electronics engineering from the University of Pisa, Italy, in 1972.

From 1975 to 1978 he was Associate Professor of Electrical Engineering at the University of Pisa. From 1978 to 1979 he was on a NATO-supported Fellowship Leave in the Department of Electrical Engineering, University of Southern California, Los Angeles. In 1982 he was Visiting Associate Professor in the Department of Electrical and Computer Engineering, University of Massachusetts, Amherst. From 1976 to 1986 he was also a Research Scientist of the Italian National Research Council (CNR) at the Centro di Studio per Metodi e Dispositivi di Radiotrasmissione, Pisa. Since 1986 he has been Professor of Electrical Engineering, University of Genoa, Italy. His research interests are in laser and RF digital communication systems and digital signal processing in communications.