Electronics and Communications in Japan, Part 3, Vol. 87, No. 3, 2004
Translated from Denshi Joho Tsushin Gakkai Ronbunshi, Vol. J86-A, No. 5, May 2003, pp. 517–528

Constrained Blind Equalizer Using Adaptive Nonlinear Filter

Ryo Koizumi,[1] Teruyuki Miyajima,[2] and Kazuo Yamanaka[2]

[1]Graduate School of Science and Engineering, Ibaraki University, Hitachi, 316-8511 Japan

[2]Faculty of Engineering, Ibaraki University, Hitachi, 316-8511 Japan

SUMMARY

In this paper, a new adaptive equalizer that adjusts the nonlinear filter without training signals is proposed and its performance is evaluated. In the proposed equalizer, the nonlinear filter has constraints, while learning is carried out by means of the output energy minimization algorithm. Since the proposed equalizer has nonlinearity, it has a better bit error rate than the linear equalizer. Further, since no training signal is needed, it is to be expected that the transmission efficiency is superior to that in an equalizer with training signals. In order to evaluate the performance limitations of the proposed equalizer, an optimum equalizer with constrained filter input is derived. By means of simulations, a performance evaluation is carried out and the effectiveness of the proposed equalizer is demonstrated. Further, in order to improve the performance of the proposed equalizer for a linear communication channel, an intersymbol interference canceler and an algorithm for correcting the error are proposed and their effectiveness is demonstrated. Electron Comm Jpn Pt 3, 87(3): 17–29, 2004; Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/ecjc.10145

Key words: intersymbol interference; blind equalization; nonlinear adaptive filter; output energy minimization; neural network.

1. Introduction

When high-speed data transmission is carried out, intersymbol interference along the propagation channel is an important problem. Equalizers are used for suppressing the effect of this intersymbol interference and estimating accurate transmitted signals. It is known that an equalizer with a nonlinear structure has better performance than a linear equalizer in a linear communication channel to which white Gaussian noise is added [1]. To date, nonlinear adaptive equalizers using nonlinear filters such as the layered neural network [2], Volterra filter [3], and RBF (Radial Basis Function) network [4] have been investigated. These nonlinear adaptive equalizers adaptively adjust the weights in a receiver by means of known training signals. However, since many training signals are needed before convergence, the transmission efficiency is found to be low.

In order to resolve this low efficiency by the use of these training signals in linear adaptive equalizers, blind equalizers have been studied, in which the linear filter can be adjusted without training signals. Typical methods include the CMA (Constant Modulus Algorithm) [5] and the MOE (Minimum Output Energy) algorithm minimizing the output energy of the constrained linear filter [6]. These linear blind equalizers have high transmission efficiency since no training signals are used.

If on the other hand nonlinear equalizers can be adjusted by a blind process, even better performance may be expected. To date, research on such nonlinear blind equalizers has included studies based on learning of the layered neural network approximated by several FIR filters by means of CMA [7] and by the RRBF (Recurrent RBF) [8]. Sufficient discussions have not yet been carried out.

In this paper, a new nonlinear blind equalizer is proposed. In the proposed equalizer, the nonlinear filter is trained in such a way that the output energy is minimized. The proposed configuration can be thought of as a replacement of the linear filter in the equalizer in Ref. 6 with a nonlinear filter. Since the proposed equalizer has nonlinearity, better performance than in a linear equalizer is expected. Since no training signals are needed, the transmission efficiency is expected to be superior to that in an equalizer using training signals. When the performance limit of the equalizer is discussed, an optimum equalizer based on the known transmission characteristics can serve as a reference. This is known as a Bayesian equalizer [2]. In this paper, a constrained optimum equalizer is derived. With reference to this configuration, the performance of the proposed equalizer is discussed. In the present paper, a method is also discussed for eliminating the intersymbol interference and correcting the decision error in order to improve the performance of the proposed equalizer.

2. Communication Model

2.1. Linear communication channel and adaptive equalizer

The communication model dealt with in this paper is shown in Fig. 1. Here, d_n denotes the transmitted data and consists of an independent sequence with +1 and −1 taken at equal probabilities; w_n is white Gaussian noise with a mean of 0 and a variance of σ^2 and is assumed independent of the transmitted signal; r_n is the channel output, u_n is the observed received signal, and y_n is the equalizer output. The received signal is given by

$$u_n = c_1 d_n + \sum_{i=2}^{L} c_i d_{n-i+1} + w_n \tag{1}$$

where c_i denotes the strength of each arriving wave. The first term on the right-hand side of Eq. (1) is the direct wave component and the second term expresses the delayed components. The equalizer estimates the transmitted signal d_n by means of the received signal vector $\mathbf{u}_n = [u_n, u_{n-1}, \ldots, u_{n-M+1}]^T$. The adaptive equalizer has adjustable coefficients in the filter so that the filter coefficients can be adjusted adaptively in response to the communication channel.

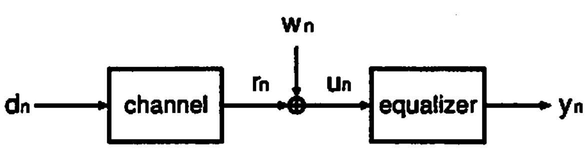

Fig. 1. Communication model.

2.2. Conventional nonlinear adaptive equalizer

The conventional nonlinear adaptive equalizer that needs training signals has the configuration shown in Fig. 2. Its output is as follows:

$$y_n = F(\mathbf{u}_n) \tag{2}$$

where F denotes a nonlinear adaptive filter with adjustable coefficients.

The data decision value can be obtained as $Y_n = \mathrm{sgn}(y_n)$. In this equalizer, during the learning period, the nonlinear filter F is adjusted in such a way that the mean square error given by the following equation is reduced by using the known training data d_n:

$$J = E[(y_n - d_n)^2] \tag{3}$$

As an example, if a layered neural network is used as a nonlinear filter, learning is carried out by the error back-propagation method [2].

When the transmitted signal is estimated from the received signal vector, the ensemble of the received signal vector, as the boundary between the received signal region with decision of −1 and that with a decision of +1 is called the decision boundary. As seen in Section 4, the decision boundary of an optimum equalizer is in general a complex curved hypersurface. In contrast to the linear equalizer for which the decision boundary is a hyperplane, the decision boundary for a nonlinear equalizer can be an arbitrary curved hypersurface. Hence, the nonlinear adaptive equalizer exhibits better performance than the linear adaptive

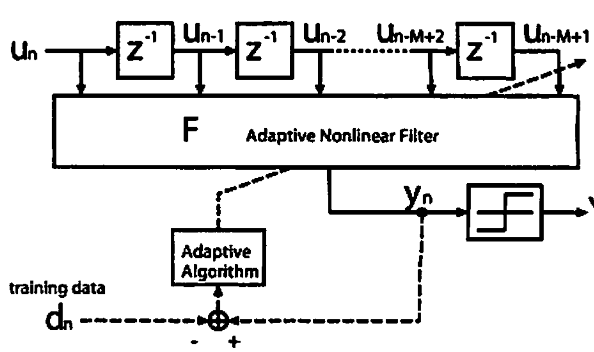

Fig. 2. Nonlinear adaptive equalizer.

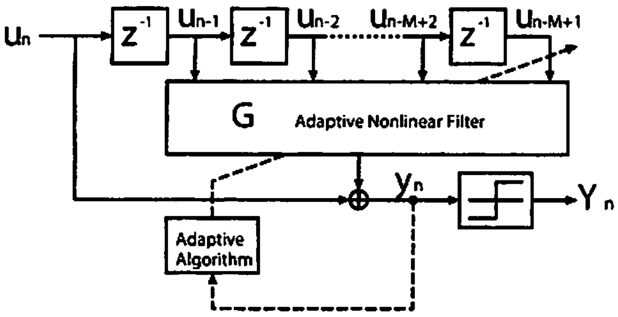

Fig. 3. Constrained nonlinear blind equalizer.

equalizer [2]. However, since training signals are used in the conventional nonlinear adaptive equalizer, transmission of the information data is not possible during the learning period. Hence, the transmission efficiency is low.

3. Proposal of Constrained Nonlinear Blind Equalizer

A configuration of the constrained nonlinear blind equalizer proposed in this paper is shown in Fig. 3. Here, G denotes a nonlinear adaptive filter with adjustable coefficients. As a nonlinear filter, a layered neural network, an RBF network, and a Volterra filter can be used. In the constrained nonlinear blind equalizer, the latest received signal u_n in the received signal vector $\mathbf{u}_n = [u_n, u_{n-1}, \ldots, u_{n-M+1}]^T$ incident on the filter is retained (this is called the constraint for the nonlinear filter), while the remaining received vector $\mathbf{v}_n = [u_{n-1}, u_{n-2}, \ldots, u_{n-M+1}]^T$ is input to the filter. The equalizer output is given by

$$y_n = u_n + G(\mathbf{v}_n) \tag{4}$$

where G is assumed to take only one value for a given $\mathbf{v}_n$. The operating principle of this equalizer is as follows. If d_n is the desired component, u_n contains both the desired component and the interference component, whereas $\mathbf{v}_n$ contains only the interference component. Therefore, only the desired component can be extracted by canceling the interference component of u_n by the interference component output from the nonlinear filter. The cost function is the mean square of the equalizer output given by

$$J = E[y_n^2] \tag{5}$$

The coefficients of G are updated in such a way that the above quantity is reduced. Since the desired component d_n is contained only in u_n and has no correlation with the interference component (delayed component), it is to be expected that the magnitude of the interference component can be kept small by minimizing the output energy while the desired component is maintained. Since the proposed equalizer has nonlinearity, it is to be expected that its performance is better than that of a linear equalizer. On the other hand, because of the constraint, there is a possibility that a decision boundary of an arbitrary curve may not be formed. Hence, in Section 4 we derive a constrained optimum equalizer, which is compared with an optimum equalizer without constraint, in order to study the performance limitations of the proposed equalizer.

4. The Optimum Equalizer

4.1. The Bayesian equalizer

An optimum equalizer in a nonconstrained nonlinear equalizer is given by a Bayesian equalizer. The output of a Bayesian equalizer is given by the following [2]:

$$y_n = \mathrm{sgn}(f^+(\mathbf{u}_n) - f^-(\mathbf{u}_n)) \tag{6}$$

where $f^+(\mathbf{u}_n)$ is the probability density function of $\mathbf{u}_n$ for $d_n = +1$, and $f^-(\mathbf{u}_n)$ is that of $\mathbf{u}_n$ for $d_n = -1$. The ensemble of $\mathbf{u}_n$ satisfying $f^+(\mathbf{u}_n) = f^-(\mathbf{u}_n)$ becomes the optimum decision boundary.

4.2. The constrained optimum equalizer

The decision boundary is an ensemble of the received signal vectors such that the output of the equalizer is 0. It is seen from Eq. (4) that in a constrained equalizer there is only one u_n,

$$u_n = -G(\mathbf{v}_n) \tag{7}$$

that satisfies the output $y_n = 0$ for the past received signal vector $\mathbf{v}_n$. Under this condition, the optimum equalizer is derived as follows:

$$y_n = \mathrm{sgn}(u_n - a_0) \tag{8}$$

$$a_0 = \arg\min_a \left\{ \int_{-\infty}^{a} f^+(u_n|\mathbf{v}_n)du_n + \int_a^{\infty} f^-(u_n|\mathbf{v}_n)du_n \right\} \tag{9}$$

Here, the ensemble $\mathbf{u}_n = [u_n, \mathbf{v}_n^T]^T$ satisfying $u_n = a_0$ (or $y_n = 0$) is the optimum decision boundary. The quantity within the medium brackets on the right-hand side of Eq. (9) is the error rate for the decision boundary of a and hence a_0 is chosen in such a way that this is minimized. The first term in the medium brackets is the error rate for $d_n = +1$ and the second term is that for $d_n = -1$.

4.3. Decision boundary of an optimum equalizer

The decision boundaries of the Bayesian equalizer and the constrained optimum equalizer derived by the above method are presented in Figs. 4 to 7. The parameters are given in Table 1. Here, SNR = 7 dB. The following two are used as a model for the communication channel:

$$\begin{aligned} &\textit{channel } 1 \quad u_n = d_n + 0.5d_{n-1} + w_n \\ &\textit{channel } 2 \quad u_n = d_n + 1.5d_{n-1} + w_n \end{aligned}$$

Here, *channel* 1 is the minimum phase channel and *channel* 2 is an example of a linear nonminimum phase channel. In the absence of noise, the received signal vector $\mathbf{r} = [r_n, r_{n-1}]^T$ is distributed as indicated by "+" and "×" in Figs. 4 through 7 due to intersymbol interference. Here, "+" denotes the plot of the received signal vector for $d_n = +1$ while "×" denotes that for $d_n = -1$. If noise is present, the actual received signals are dispersed, depending on the dispersion of the noise centered around "+" and "×." The problem of equalization can be treated as the problem of dividing the received signal plane by a boundary line into the region to be judged as +1 and that to be judged as −1. When the optimum decision boundaries derived here are compared, it is to be expected from Figs. 4 and 5 that the constrained optimum equalizer provides a curve similar to that for Bayesian equalizer for *channel* 1 and that both equalizers have similar performance. Next, from Figs. 6 and 7, there appears a difference in the decision boundary between Bayesian equalizer and the constrained optimum equalizer in the case of *channel* 2. In both cases, the received signal points can be accurately divided in the absence of noise. It is considered that the Bayesian equalizer is more robust to the noise.

Table 1. Parameters

Communication channel	Linear (FIR)
Number of arriving waves L	2
Number of elements M in the equalizer input $\mathbf{u}_n$	2
Desired signal	d_n
Definition of SNR	SNR [dB] = $10 \log \frac{1}{2} (c_1^2/\sigma^2)$

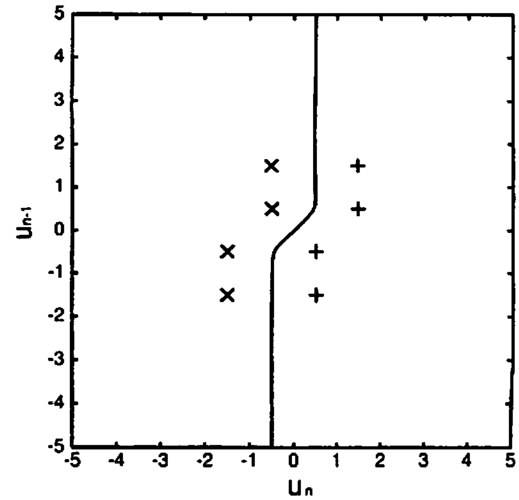

Fig. 4. Bayesian decision boundary for *channel* 1.

4.4. BER characteristics of the optimum equalizer

The bit error rate (BER) of the optimum equalizer is derived by computer simulations. The conditions are identical to those in Table 1. The communication channels are *channel* 1 and *channel* 2. The BER characteristics for *channel* 1 are given in Fig. 8 and those for *channel* 2 in Fig. 9. In *channel* 2, for which there is a difference in the decision boundaries above, the BER characteristics of the constrained optimum equalizer become degraded. Hence,

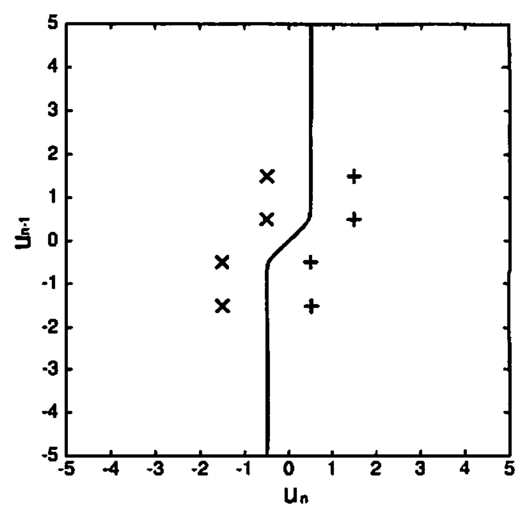

Fig. 5. Decision boundary by constrained optimal equalizer for *channel* 1.

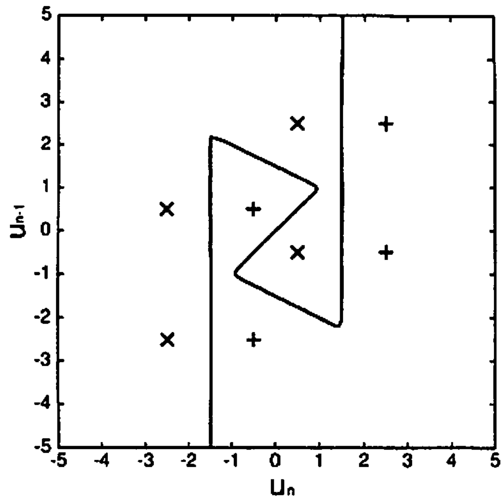

Fig. 6. Bayesian decision boundary for *channel* 2.

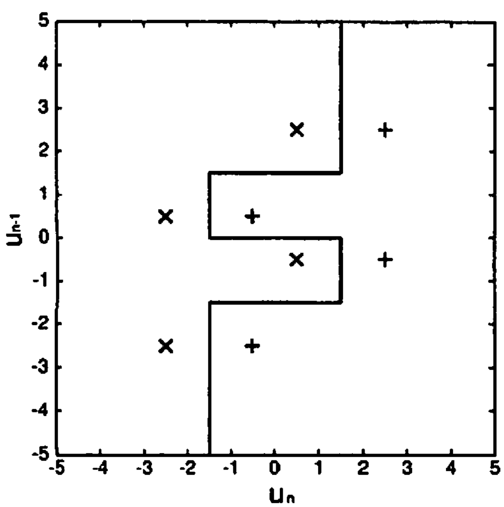

Fig. 7. Decision boundary by constrained optimal equalizer for *channel* 2.

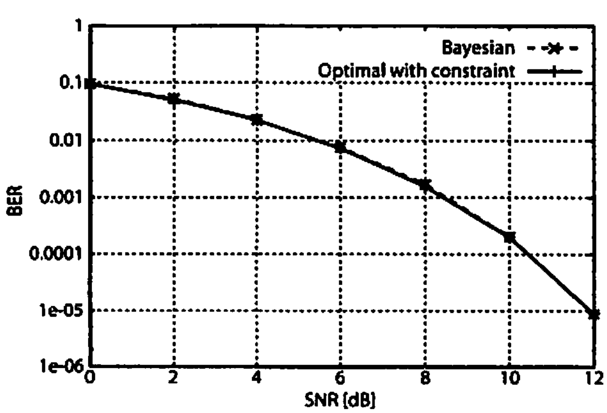

Fig. 8. BER curves by Bayesian equalizer and constrained optimal equalizer for *channel* 1.

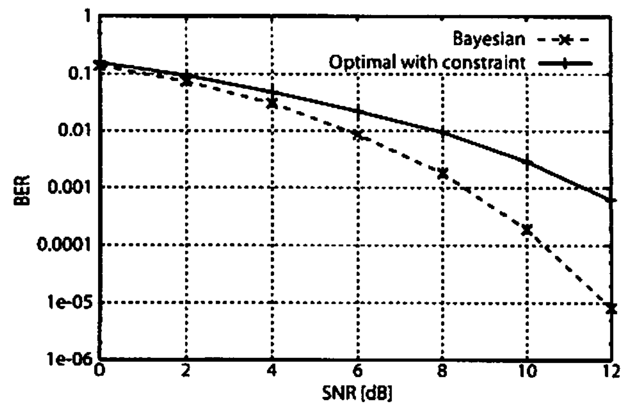

Fig. 9. BER curves by Bayesian equalizer and constrained optimal equalizer for *channel* 2.

it is found that there are cases in which the proposed constrained nonlinear blind equalizer is inferior in terms of performance limitation to the conventional nonlinear adaptive equalizer. Nevertheless, it is confirmed that correct division of the received signal points by the proposed scheme is possible in the absence of noise in both minimum phase and nonminimum phase communication channels.

In general, differential coding must be used in the blind equalizer because ambiguity between positive and negative values remains in the output. By using differential coding, the BER characteristics are somewhat degraded. This is specifically described in the next section.

5. Performance Evaluation

5.1. Computer simulation

A performance evaluation of a constrained nonlinear blind equalizer was performed by computer simulation. For comparison, simulations have also been carried out for the nonlinear adaptive equalizer using training signals as proposed in Ref. 2 and the linear equalizer for an FIR filter with two taps. In the following, the nonlinear adaptive equalizer using the training signals proposed in Ref. 2 is called the "conventional nonlinear equalizer." The layered neural network is used as a nonlinear filter. The conventional nonlinear equalizer is trained by the error back-propagation method. The linear equalizer performs learning by the LMS algorithm.

The output of the proposed equalizer is as follows when a layered neural network is used:

$$y_n = u_n + G(\mathbf{v}_n)$$

$$= u_n + \sum_{i=1}^{I} q_i f \left[\sum_{m=1}^{M-1} p_{im} u_{n-m} + p_{i0} \right] \quad (10)$$

Here, $M - 1$ is the number of inputs into the layered neural network and I is the number of neurons in the intermediate layer. The output $f[]$ of the intermediate layer neurons is a sigmoid function and is given by

$$f[x] = \frac{1 - e^{-x}}{1 + e^{-x}} \quad (11)$$

where p_{im} ($m = 1, 2, \ldots, M - 1$) is the weighting coefficient for the connection between the input layer and the intermediate layer, p_{i0} is the threshold value for the intermediate layer, and q_i is the weighting coefficient for the connection between the intermediate layer and the output layer. The updating of the weight coefficient is adjusted in such a way as to minimize the instantaneous approximation of the mean square of the output in terms of the square of the outputs at each time by means of the stochastic gradient method. The specific updating equation is given below. The sum of the inputs into the i-th intermediate neuron is expressed as

$$e_{in} = \sum_{m=1}^{M-1} p_{im} u_{n-m} + p_{i0} \quad (12)$$

Further, the output from the i-th intermediate layer neuron is expressed by

$$f_{in} = f\left[e_{in}\right] \quad (13)$$

Thus, the updating equations for the weighting coefficient p_{im} for the connection between the input layer and the intermediate layer and the threshold value p_{i0} are given by

$$p_{im(n+1)} = p_{imn} + \Delta_{imn} \quad (14)$$

$$\Delta_{imn} = -A \frac{\partial y_n^2}{\partial p_{imn}} + B \Delta_{im(n-1)} \quad (15)$$

$$\begin{aligned} \frac{\partial y_n^2}{\partial p_{imn}} &= 2 y_n q_{in} (1 - f_{in})(1 + f_{in}) u_{n-m} \\ &\quad (m = 1, 2, \ldots, M - 1) \\ &= 2 y_n q_{in} (1 - f_{in})(1 + f_{in}) \\ &\quad (m = 0) \end{aligned} \quad (16)$$

where p_{imn} is the weight coefficient for connection of the input layer and the intermediate layer at time n, Δ_{imn} is the size of the update for the weight coefficient, A is the step size parameter determining the magnitude of one update, and B is the moment parameter. The updating equation for the weight coefficient q_i for connection between the intermediate layer and the output layer is

$$q_{i(n+1)} = q_{in} + \Delta_{in} \quad (17)$$

$$\Delta_{in} = -A \frac{\partial y_n^2}{\partial q_{in}} + B \Delta_{i(n-1)} \quad (18)$$

$$\frac{\partial y_n^2}{\partial q_{in}} = 2 y_n f_{in} \quad (19)$$

where q_{in} is the weight coefficient for connection between the intermediate layer and the output layer at time n, and Δ_{in} is the size of the update of the weight coefficient. The specifications of the subsequent simulation are shown in Table 2. The step size parameter and the moment parameter are selected in such a way that the convergence speed is faster among the values minimizing the error rate. The number of intermediate neurons I is 4 in the layered neural network.

As mentioned in Section 4, differential coding is actually needed in the proposed nonlinear blind equalizer. On the other hand, the conventional nonlinear equalizer does not need differential coding. In the proposed nonlinear blind equalizer, performance degradation may occur due to blinding and to use of differential coding. To clarify these issues, the results for the proposed nonlinear blind equalizer are presented with and without differential coding in Sections 5.3 and 5.4. No differential coding is used in the linear equalizer presented in Section 5.3.

In Sections 5.3 and 5.4, simulations are also carried out while varying c_2 randomly for each trial in addition to the fixed communication channels such as *channels* 1 and 2. Specifically, the following communication channels are considered:

$$\begin{aligned} &\text{channel } A \quad u_n = d_n + (0.5 + rand) d_{n-1} + w_n \\ &\text{channel } B \quad u_n = d_n + (1.5 + rand) d_{n-1} + w_n \end{aligned}$$

Table 2. Simulation parameters

Communication channel	Linear (FIR)
Number of arriving waves L	2
Number of elements M in the equalizer input $\mathbf{u}_n$	2
Desired signal	d_n
Step size A	0.001
Step size B	0.9
Definition of SNR	SNR [dB] = $10 \log \frac{1}{2} (c_1^2/\sigma^2)$

where *rand* is a uniform random value between −0.2 and +0.2: *channel A* is a minimum phase channel and *channel B* is a nonminimum phase channel.

5.2. Learning convergence

The learning convergence of the proposed equalizer for *channel* 1 and *channel* 2 is shown in Fig. 10. The horizontal axis indicates the number of iterations and the vertical axis is the BER. The SNR is 7 dB. Learning is carried out once each time 1 bit of signal is received. After 100 learning operations, the weight coefficients are fixed and the BER is derived for 30,000 bits of received signal. Simulations are carried out for 100 trials while varying the initial values of the filter coefficients, transmitted signals, and noise. The averages of the results are obtained. It is found that the constrained nonlinear blind equalizer using the layered neural network can reduce the BER for *channel* 1 and *channel* 2 so that learning converges. Hence, learning by output energy minimization is confirmed to be effective.

5.3. Evaluation of BER characteristics

The BER characteristics are obtained for a linear equalizer after 1000 learning operations (Linear), the conventional nonlinear equalizer after 40,000 learning operations (Nonlinear), the constrained nonlinear blind equalizer after 40,000 learning operations (Nonlinear Blind), and the constrained nonlinear blind equalizer after 40,000 learning operations with differential coding (Nonlinear Blind D). The BER is obtained for 50,000 bits of received signal after learning. In the simulations, the initial values of the filter coefficients, transmitted signals, and noise are varied for 100 trials and the averages of the results are taken. The results are shown in Figs. 11 and 12. The numbers and letters in brackets in the figures indicate the communication channels. In Fig. 11, the BER characteristics are compared for *channel* 1 and *channel A*. The proposed equalizer has the same performance as the conventional nonlinear equalizer. As described in Section 4, both equalizers can form the

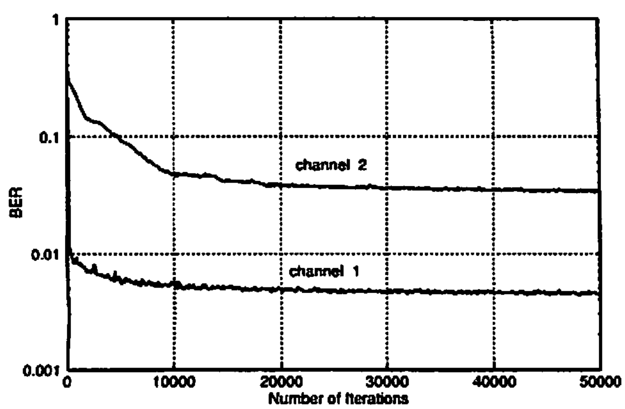

Fig. 10. BER of constrained nonlinear blind equalizer.

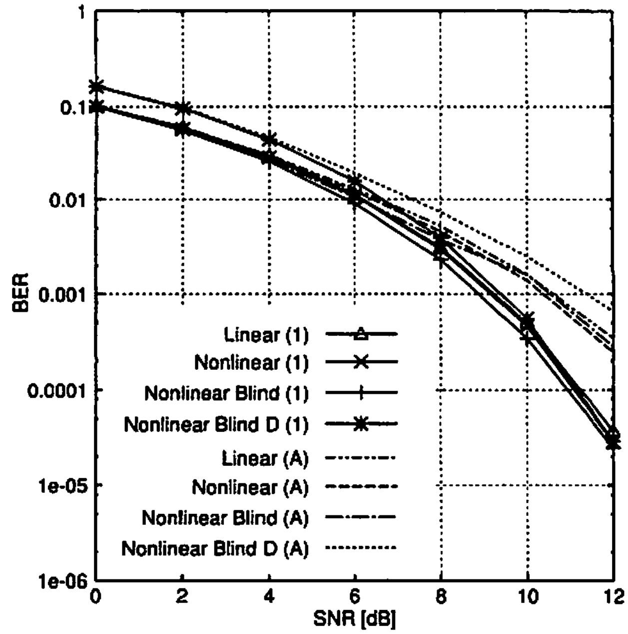

Fig. 11. BER comparison for *channel* 1, *channel A*.

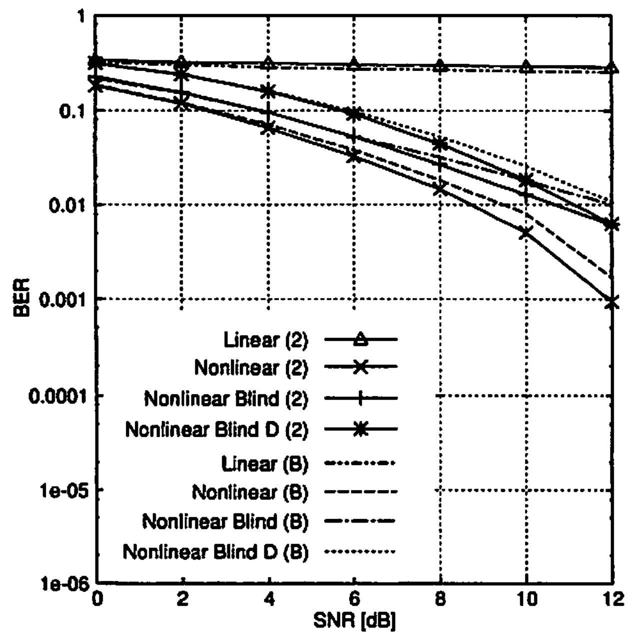

Fig. 12. BER comparison for *channel* 2, *channel B*.

same decision boundary for the minimum phase channel. Also, it is confirmed that the performance degradation due to differential coding is slight. Figure 12 presents a comparison of the BER characteristics in *channel* 2 and *channel B*. Here, it is found that the BER characteristics of the constrained nonlinear blind equalizer are worse than those of the conventional nonlinear equalizer. As described in Section 4, this is performance degradation due to constraint. Even here, the performance degradation due to differential coding is slight. On the other hand, the decision boundary of the linear equalizer is a hyperplane and hence the BER characteristics are extremely poor, since the signal points cannot be divided correctly even without noise.

The results for the case in which a communication channel is given randomly in all equalizers indicate worse BER characteristics than those for a fixed channel. When the channel is given randomly, the receiving signal points without noise in the receiving signal space are very close if the intensity c_2 of the delayed wave approaches the one c_1 of the direct wave. This is considered to occur because the equalizer inputs are limited to two. It is confirmed by simulation that the degradation of the BER characteristics can be reduced somewhat by increasing the number of equalizer inputs.

Let us compare the BER characteristics for *channel* 1 and *channel* 2 in Figs. 8 and 11, 9 and 12. The characteristics of the nonlinear equalizer after learning do not reach those of the optimum equalizer, representing the performance limit. This implies that the weight coefficients of the layered neural network are not adjusted optimally even if learning converges. No improvement is seen when similar simulations are carried out with an increase in the number of intermediate layer neurons in the layered neural network. The causes may be falling into local minima and variations of the weight coefficients after convergence due to noise. This depends on the nonlinear filter and the learning method. This issue merits future study.

Let us describe the behavior of the proposed equalizer in a communication channel in which the direct wave component c_1d_n vanishes, as in the raised cosine channel [9]. When the direct wave component c_1d_n vanishes, the transmitted signals contained in the latest received signal u_n are $d_{n-1}, d_{n-2}, \ldots, d_{n-L+1}$. Since $d_{n-2}, d_{n-3}, \ldots,$ $d_{n-L-M+2}$ are contained in the nonlinear filter input $\mathbf{v}_n$, it is to be expected that only d_{n-1} will remain if u_n is fixed. Actually a simulation was carried out to confirm that the proposed equalizer acts to estimate the transmitted signal d_{n-1} one time step past in this communication environment.

In this paper, $M = 2$ is used so that the discussion in terms of the formation capability of the decision boundary in the input signal space can be understood intuitively. If M is increased in a linear equalizer, the performance is improved somewhat for *channel* 1, but it is confirmed that no performance improvement takes place for *channel* 2. In each case, the proposed equalizer is confirmed to have BER characteristics superior to the linear equalizer for the same M.

5.4. Evaluation of transmission efficiency

When the channel is time-varying, it is necessary to perform relearning frequently. Thus, if the training signal is used, the ratio of the number of training data to the number of transmitted data becomes increased so that the number of information data that can be transmitted per unit time is reduced. No such situation occurs *in the blind* equalizer because no learning is carried out. Hence, this "transmission efficiency," defined as the number of information data that can be transmitted within a unit time, is evaluated. Let the number of transmitted data be D, the number of training data contained in the transmitted data be T, and the number of information data be I. Note that $D = I$ in the blind equalizer and $D = I + T$ in an adaptive equalizer that requires training signals. The transmission efficiency ρ is defined as

$$\rho = \frac{I_s}{D} \tag{20}$$

where I_s is the number of information data that are correctly demodulated from the information data I. Computer simulations are carried out to derive the transmission efficiency, assuming relearning for every 500 bits in a constrained nonlinear blind equalizer and a conventional nonlinear equalizer. The parameters are listed in Table 2.

In order to first derive the optimum number of training data in the re-learning per 500 bits, the relationship between the number of training data T and the transmission efficiency ρ is studied. In the case of SNR = 7 dB, the optimum number of training data is derived. In the conventional nonlinear equalizer, T bits of training data in the 500 bits of received data are repeated offline for learning. After learning, the remaining 500-T bits are demodulated. Since the proposed equalizer does not require a training signal, all of the 500 bits of received data are used repeatedly for learning regardless of T, and 500 bits are used for demodulation. In the simulation, the initial values of the filter coefficients, transmitted signals, and noise are varied and averages are taken after 500 trials. The results are shown in Figs. 13 and 14. It is found that the transmission efficiency of the conventional nonlinear equalizer does not compete with that of the proposed equalizer with and without differential coding for any number of training data. When the results of the conventional nonlinear equalizer are observed, the transmission efficiency is found to be maximum at $T = 20$ in the minimum phase channel and at $T = 30$ in the nonminimum phase channel. Since it is not known a

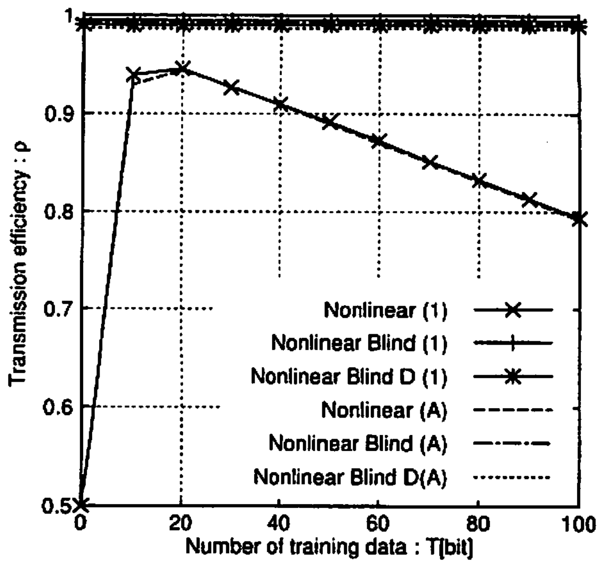

Fig. 13. Effect of the number of training data to transmission efficiency (*channel* 1, *channel A*).

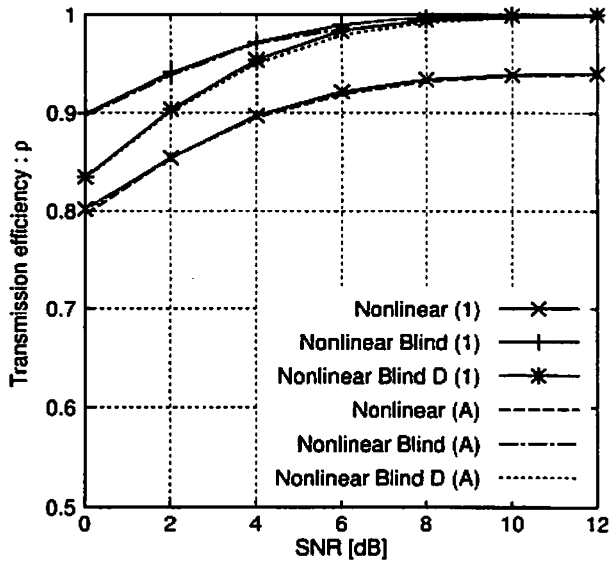

Fig. 15. Comparison of transmission efficiency (*channel* 1, *channel A*).

priori whether the channel is a minimum phase shift system or a nonminimum phase shift system, the number of training data needed to obtain a good transmission efficiency is assumed to be $T = 30$. This is considered to be the optimum number of training data for relearning per 500 bits. Here, the optimum number of training data is derived for SNR = 7 dB. In the case of SNR = 0, 2, 4, 6, 8, and 12 dB, it is confirmed through simulation that the optimum number of training data is also $T = 30$.

Next, the relationship between the SNR and the transmission efficiency is studied for the optimum number of training data $T = 30$ obtained in the above simulation. The results are shown in Figs. 15 and 16. The results confirm that there is little difference in transmission efficiency between the fixed communication channel and the random one. Also, regardless of whether or not differential coding is used, the proposed equalizer has a transmission efficiency better than that of the conventional nonlinear equalizer.

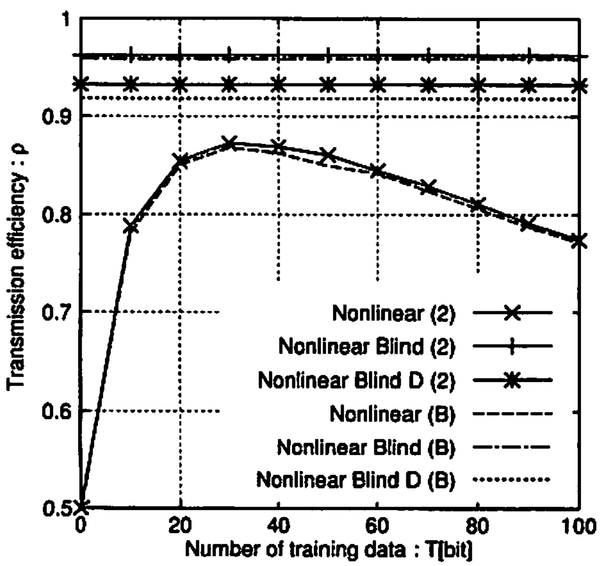

Fig. 14. Effect of the number of training data to transmission efficiency (*channel* 2, *channel B*).

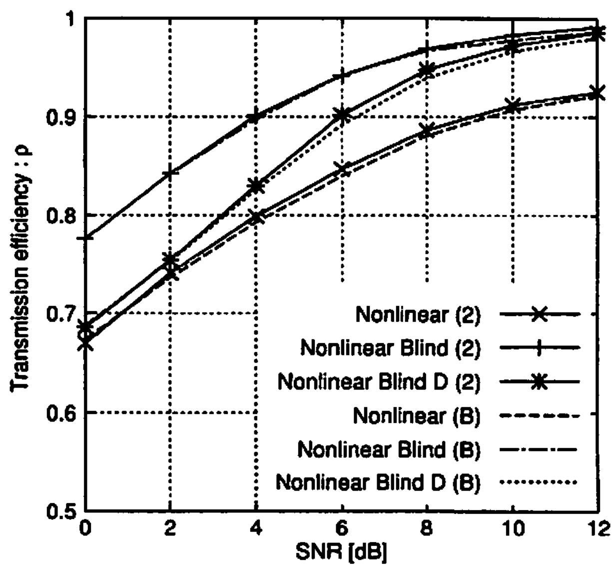

Fig. 16. Comparison of transmission efficiency (*channel* 2, *channel B*).

The following is confirmed by the above performance evaluation. Since the proposed equalizer can form a decision boundary identical to the one in the conventional nonlinear equalizer in the minimum phase channel, it has the same BER characteristics as the conventional nonlinear equalizer. On the other hand, the proposed equalizer has BER characteristics in the nonminimum phase channel that are inferior to the conventional nonlinear equalizer due to the difference in the decision boundary caused by the constraint. However, in terms of transmission efficiency, the proposed scheme is superior to the conventional nonlinear equalizer in both the minimum phase and nonminimum phase channels within the range of the simulations discussed here.

6. Improvement of Constrained Nonlinear Blind Equalizer

As described above, there are cases in which sufficient performance cannot be obtained by the proposed equalizer due to constraint. Let us therefore study the performance improvement of the proposed equalizer for a linear communication channel. We propose to construct an intersymbol interference cancellation system by attaching a linear adaptive filter as a postprocessing system to a constrained nonlinear blind equalizer. Further, an algorithm is proposed that nullifies the error caused by the propagation of the error from the equalizer to the linear filter, and its effectiveness is demonstrated.

6.1. Intersymbol interference cancellation system

The configuration of the system proposed here is shown in Fig. 17. The portion enclosed by dotted lines on the lower right is the intersymbol interference canceler. An estimated value sequence $Y_{n-i+1} (i = 2, 3, \ldots, N)$ of past

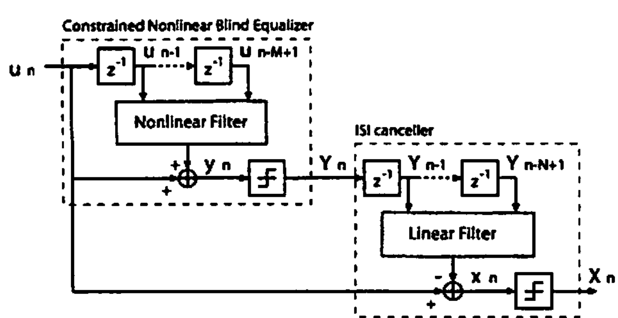

Fig. 17. Intersymbol interference cancel system.

transmitted signals determined by the constrained nonlinear blind equalizer is incident on the linear filter. The output of the intersymbol interference is

$$
\begin{aligned}
x_n &= u_n - \sum_{i=2}^{N} h_i Y_{n-i+1} \\
&= \sum_{i=2}^{L} c_i d_{n-i+1} - \sum_{i=2}^{N} h_i Y_{n-i+1} \\
&\quad + c_1 d_n + w_n
\end{aligned}
\tag{21}
$$

where h_i is the adjustable coefficient of the linear filter. Then, if h_i is set such that $h_i = c_i (i \leq L), 0 (i > L)$ when $N \geq L$ and all decisions by the equalizer are correct, the interference components contained in u_n can be eliminated by using the interference replica generated by the linear filter.

Since c_i is unknown in practice, h_i is adjusted by an adaptive algorithm. The mean square of the output of the intersymbol interference canceler is used as the cost function:

$$
J = E[x_n^2] \tag{22}
$$

The filter coefficient h_i is adjusted in such a way that the above is made smaller. The specific updating equation is

$$
h_{i(n+1)} = h_{in} + C x_n Y_{n-i+1} \tag{23}
$$

where C is the step size parameter used to determine the step of one iteration. Assuming $N \geq L$, let $\mathbf{Y} = [Y_{n-1}, Y_{n-2}, \ldots, Y_{n-N+1}]^T$, $\mathbf{h} = [h_2, h_3, \ldots, h_N]^T$, $\mathbf{d} = [d_{n-1}, d_{n-2}, \ldots, d_{n-N+1}]^T$ and $\mathbf{c} = [c_2, c_3, \ldots, c_L, 0, \ldots, 0]^T$. When the cost function is subjected to partial differentiation with respect to $\mathbf{h}$, we let $\partial J / \partial \mathbf{h} = 0$. Then, it follows that

$$
\mathbf{h}^T E[\mathbf{Y}\mathbf{Y}^T] = \mathbf{c}^T E[\mathbf{d}\mathbf{Y}^T] \tag{24}
$$

Assuming that $\mathbf{Y} = \mathbf{d}$ and using $E[d_i d_j] = 0 (i \neq j), 1 (i = j)$, we obtain

$$
\mathbf{h} = \mathbf{c} \tag{25}
$$

Hence, under the assumption that $\mathbf{Y} = \mathbf{d}$, $N \geq L$, the filter coefficient vector minimizing the cost function is equal to $\mathbf{c}$, and the impulse response of the communication channel is derived. However, since the error rate of the equalizer decision is not 0 in general, $\mathbf{Y} = \mathbf{d}$ does not necessarily hold. The filter coefficient $\mathbf{h}$ trained by the output energy minimization is somewhat deviated from $\mathbf{c}$.

6.2. Confirmation of operation of the intersymbol interference cancellation system

The operation of the intersymbol interference canceler was confirmed by computer simulation. The communication channels used were *channel* 1 and *channel* 2. In addition, SNR = 7 dB, the number of inputs to the canceler was $N = 2$ (so that the number of inputs to the linear filter was 1), and the step size parameter C was 0.001. The changes of the filter coefficients for 50,000 updates are shown in Fig. 18. From the results, it is found that h approaches 0.5 for *channel* 1 ($c_2 = 0.5$) and almost 1.5 for *channel* 2 ($c_2 = 1.5$), so that the operation is correct. If an average of h_2 after convergence is taken, the result is 0.4978 for *channel* 1 and 0.1414 for *channel* 2. In particular, for *channel* 2, the error rate of the decision Y_n of the equalizer is high so that h_2 deviates from the target value of 1.5. In general, the number of arriving waves L is unknown. As described in Section 6.1, this situation can be dealt with if N is chosen with a margin. Under the same conditions as above, the simulation was repeated after N was increased. Then, h_3 and higher-order terms converged to 0, indicating that the operation is correct even if N is increased.

6.3. Error correcting algorithm

When an error occurs in the decision Y_k of the equalizer at time k in the intersymbol interference cancellation system, a false input is incident on the linear filter after time $k + 1$. Hence, error is more likely in the decisions X_{k+1} through X_{k+N-1} in the canceler. This problem is resolved by using several characteristics of the sequences of Y_n and X_n. Since both Y_n and X_n are estimates for d_n, either Y_n or X_n is in error if the signs of Y_n and X_n are different. Since an error is more likely in X_n one time period after an error occurs in Y_n, there is a greater possibility that an error in X_n occurs after an error in Y_n. On the other hand, X_n is not affected by any error in the decision Y_k of the equalizer at the same time. From these characteristics, by comparing the signs of Y_n and X_n sequentially in time, it is possible to infer that the difference in sign at the first occurrence is due to an error in Y_n, whereas that from a time one time step later through $N - 1$ time steps later is caused by an error in X_n. An algorithm can be derived in which an error in Y_n is detected by comparison of signs using the above reasoning and the error is corrected. This algorithm is given in Eqs. (26), (27), and (28). In order to use the exclusive-OR, Y_n and X_n taking values of ± 1 are replaced with the values of 0 and 1 by using $\bar{Y}_n = (Y_n + 1)/2$ and $\bar{X}_n = (X_n + 1)/2$:

$$\gamma_n = \tilde{X}_n \oplus \tilde{Y}_n \tag{26}$$

$$\varepsilon_n = \gamma_n \prod_{k=1}^{N-1} (1 - \gamma_{n-k}) \tag{27}$$

$$O_n = 2(\tilde{Y}_n \oplus \varepsilon_n) - 1 \tag{28}$$

where $\oplus$ is the operator indicating the exclusive-OR and γ_n is the error detection parameter, indicating that there always exists an error in either Y_n or X_n if its value is 1. The error correcting parameter ε_n detects the error in Y_n from γ_n. The final decision modified by the error correcting parameter is O_n.

Here, only the algorithm used to detect only the error in Y_n from γ_n and to correct the decision in Y_n is presented. This process is equivalent to detection of only the error in X_n and correction of the decision in X_n. The same final decision O_n is obtained. Hence, the error correcting algorithm is one that nullifies the decision error by the intersymbol interference canceler caused by input of a false decision by the equalizer.

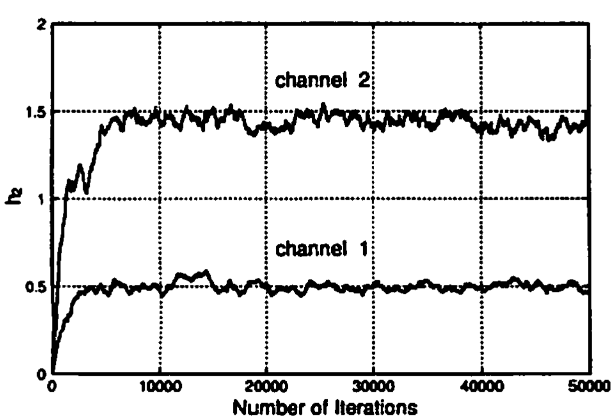

Fig. 18. Learning curves for ISI canceler.

6.4. BER characteristics after improvement

Figure 19 shows the entire structure of the improved system. Both the constrained nonlinear blind equalizer and the intersymbol interference canceler within the system are capable of learning by the output energy minimization algorithm. Therefore, blind learning is possible, yielding an improved system. By computer simulation, the decision Y_n by the constrained nonlinear blind equalizer, the decision X_n by the intersymbol interference canceler, and the final decision O_n were evaluated in terms of their BER characteristics. Updating of the filter coefficient was carried out once every time 1 bit was received. In the improved system after 40,000 updates, the BER was derived for 50,000 bits

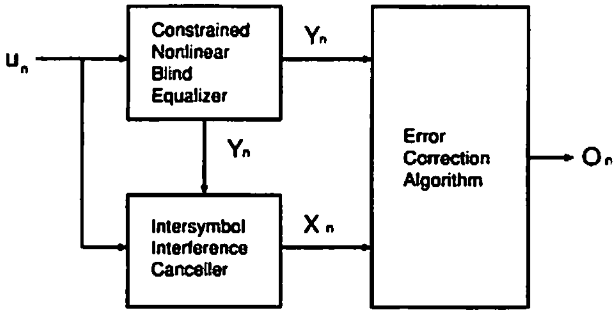

Fig. 19. Structure of improved system.

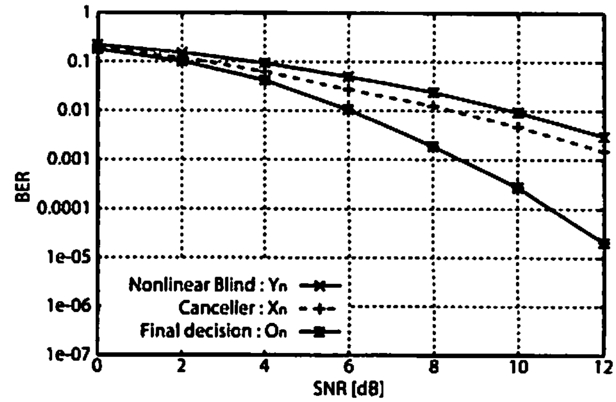

Fig. 21. BER curves for improved constrained nonlinear blind equalizer for *channel* 2.

of received signal. The communication channels used were *channel* 1 and *channel* 2. The specifications are given in Table 2. The number of inputs to the intersymbol interference was $N = 2$. The step size parameter C was 0.001. Simulation was performed 5000 times while varying the initial values of the filter coefficients, transmitted signals, and noise and then the average was taken. In Figs. 20 and 21, the BER characteristics are shown for the decision Y_n of the constrained nonlinear blind equalizer, the decision X_n of the canceler, and the final decision O_n. In Figs. 20 and 21, the decision X_n of the intersymbol interference canceler is somewhat superior to the decision Y_n of the equalizer. However, due to error propagation from the equalizer, no improvement of the BER characteristics is seen. On the other hand, significant improvement of the BER characteristics is seen in the final decision O_n in which the effect of the error propagation is nullified by error correction. The characteristics obtained in the final decision are better than the BER characteristics of the constrained optimum equalizer and also are better than the BER characteristics of the Bayesian equalizer in *channel* 1. The proposed improved system operates with only the received signals and the decision of the equalizer. Hence, this system is considered to be easily applied to other equalization systems.

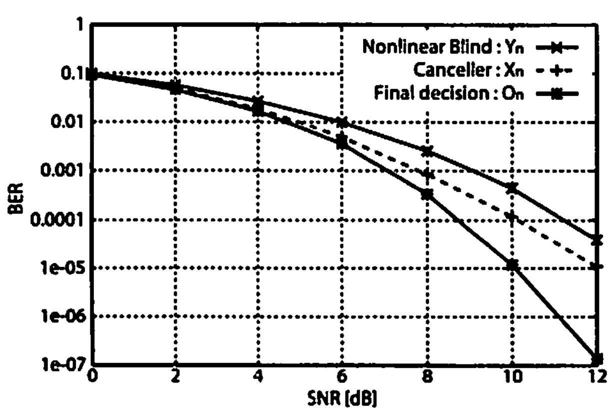

Fig. 20. BER curves for improved constrained nonlinear blind equalizer for *channel* 1.

7. Conclusions

In this paper, a new adaptive equalizer is proposed in which the nonlinear filter can be adjusted without using training signals. An optimum equalizer with a constraint in the input is derived and the performance limitation of the proposed scheme is discussed. By computer simulations, performance evaluation of the proposed system is carried out. It is shown that the proposed system is superior to the conventional equalizers in terms of transmission efficiency. Further, an enhanced system is proposed for improved performance in a linear communication channel. It is shown that the BER characteristics in the enhanced system are substantially enhanced. This enhanced system is considered applicable to other equalization systems.

In this paper, the performance of the proposed system is discussed for a linear communication channel. It is to be expected that this system will exhibit excellent characteristics for a nonlinear communication channel since it has a nonlinear characteristic. In the future, it is planned to investigate the performance for a nonlinear communication channel. Also, in the present paper, the proposed equalizer is evaluated for a time-invariant channel. In the future, evaluations for a time-varying channel in the presence of Rayleigh fading are planned.

REFERENCES

1. Ungerboeck G. Nonlinear equalization of binary signals in Gaussian noise. IEEE Trans Commun Technol 1971;19:1128–1137.
2. Gibson GJ, Siu S, Cowan CFN. The application of nonlinear structures to the reconstruction of binary signals. IEEE Trans Signal Process 1991;39:1877–1884.
3. Kajikawa Y. Present status and future prospects of adaptive Volterra filters. Trans IEICE 1999;J82-A:759–768.
4. Miyake M, Oishi K, Yamaguchi S. High speed pre-set automatic equalizer using a nonlinear function effective for RBF expansion and its evaluations. Trans IEICE 1997;J80-A:1888–1901.
5. Johnson CR, Schmiter P, Endres IJ, Behm JD, Brown DR, Casas RA. Blind equalization using the constant modulus criterion: A review. Proc IEEE 1998:86: 1927–1950.
6. Verdú S, Anderson BDO, Kennedy RA. Blind equalization without gain identification. IEEE Trans Inf Theory 1993;39:292–297.
7. Dai XH. CMA-based nonlinear blind equaliser modelled by a two-layer feedforward neural network. IEE Proc Commun 2001:148:243–248.
8. San LM, Revuelta J, Sueiro JC. Neuro Bayesian blind equalization with BER estimation in digital channels. Proc 1999 IEEE Workshop on Neural Networks for Signal Processing, p 333–342.
9. Haykin S. Adaptive filter theory, 4th ed. Prentice-Hall; 2002. Chapter 5.

AUTHORS (from left to right)

Ryo Koizumi (student member) graduated from the Department of Systems Engineering, Ibaraki University, in 2001, and completed the M.S. program in 2003. During his studies, he engaged in research on adaptive equalizers.

Teruyuki Miyajima (member) graduated from the Department of Electrical Engineering, Saitama University, in 1989, and completed the doctoral program in 1994. He then became a research associate in the Department of Systems Engineering, Ibaraki University, and has been a lecturer since 1998. He was a visiting scholar at the University of California, Davis in 2002–2003. His research interest is signal processing for communications. He received a Young Engineer Award from IEICE in 1996. He holds a doctoral degree, and is a member of IEEE.

Kazuo Yamanaka graduated from the Department of Electrical Engineering, Waseda University, in 1974, completed the doctoral program in 1979, and became a research associate at Ibaraki University. He has been a professor in the Department of Systems Engineering since 1993. He has been engaged chiefly in research on control problems in systems with uncertainty and on mathematical-physics models of neural networks. He is a member of the Institute of Electrical Engineers of Japan and the Japan Neural Network Society.