

RECURRENT CANONICAL PIECEWISE LINEAR NETWORK FOR BLIND EQUALIZATION

Xiao Liu and Tülay Adalı

Information Technology Laboratory
Department of Computer Science and Electrical Engineering
University of Maryland Baltimore County, Baltimore, MD 21250, U.S.A. *

ABSTRACT

The recurrent canonical piecewise linear (RCPL) network is applied to nonlinear blind equalization by generalizing Donoho's minimum entropy deconvolution approach. We first study the approximation ability of the canonical piecewise linear (CPL) network and the CPL based distribution learning for blind equalization. We then generalize these conclusions to the RCPL network. We show that nonlinear blind equalization can be achieved by matching the distribution of the channel input with that of the RCPL equalizer output. A new blind equalizer structure is constructed by using RCPL network and decision feedback. We discuss application of various cost functions to RCPL based equalization and present experimental results that demonstrate the successful application of RCPL network to blind equalization.

1. INTRODUCTION

Blind equalization refers to the problem of determining the transmitted symbol sequence in the presence of intersymbol interference (ISI) and noise without using a training sequence. Most of the existing blind techniques such as Bussgang algorithm, cyclic spectrum approach, and polyspectra approach, are based on the linear channel assumption. There are many cases, however, where this assumption is not true, as nonlinear devices significantly contribute to system degradation. One example is the digital satellite link [4], in which both the earth station and the satellite are equipped with amplifiers operated in a nonlinear region of the input-output characteristics for better exploitation of the power of the device. The use of the above blind techniques will suffer from a severe performance degradation for unknown nonlinear communication channels and hence the development of nonlinear blind equalization techniques carries particular significance.

In this paper, we consider application of recurrent canonical piecewise linear (RCPL) network to nonlinear blind equalization by generalizing Donoho's minimum entropy deconvolution approach [6] to the nonlinear case. We first study nonlinear filtering by canonical piecewise linear (CPL) network. Then, the corresponding conclusions are generalized for the RCPL network. We show that nonlinear blind equalization can be achieved by matching the distribution of the channel input with the distribution of the output of the RCPL equalizer. A new blind equalizer structure is constructed by using RCPL network and decision feedback. Application of various cost functions for RCPL network based equalizer is discussed and it is shown that the blind algorithm derived based on the combined Godard [7] and Vembu's convex [10] cost functions exhibits good tradeoff in terms of robustness and low equalizer prediction

error. The developed blind algorithm exhibits much faster convergence than those based on the Godard cost function or the Vembu's convex cost function alone. Simulation results are also presented to show that the RCPL network based decision feedback equalizer (RCPL-DF) outperforms both the linear decision feedback equalizer (linear-DF) and the CMA equalizer when equalizing a nonlinear channel.

2. NONLINEAR FILTERING BY CPL NETWORK

Canonical piecewise linear (CPL) network is initially introduced for nonlinear circuit analysis [5]. CPL structures provide a desirable compromise between the approximation ability of nonlinear models and the efficiency and theoretical accessibility of the linear domain, and reduce the parameter storage requirement of piecewise linear models considerably by employing a global linear representation. By using a piecewise linear filter, we can use standard linear adaptive filtering techniques to perform training tasks and can easily incorporate known statistical information into the network structure.

The CPL function is defined as [5]:

Definition 1 (*Canonical Piecewise Linear (CPL) Function*): A piecewise linear function $f: D \rightarrow Q$, with a compact subset $D \subset R^N$ and compact subset $Q \subset R^M$, is called a canonical piecewise linear (CPL) function, if it can be expressed by a *global* representation:

$$f(\mathbf{x}) = \mathbf{a} + \mathbf{B}\mathbf{x} + \sum_{i=1}^{\tau} \mathbf{c}_i |\langle \alpha_i, \mathbf{x} \rangle + \beta_i| \quad (1)$$

where $\mathbf{B} \in R^{M \times N}$, $\mathbf{a}, \mathbf{c}_i \in R^M$, $\alpha_i, \mathbf{x} \in R^N$ and $\beta_i \in R$. Based on the above definition, we show the approximation ability of CPL network by the following theorem:

Theorem 1: Let domain D be a compact space of dimension N and $\mathcal{F}$ be a set of canonical piecewise linear functions on D . Then, for any continuous function $\tilde{f}$ on D , there exists a function $f \in \mathcal{F}$ such that $|f(\mathbf{x}) - \tilde{f}(\mathbf{x})| < \epsilon$ for all $\mathbf{x} \in D$.

Proof of the theorem is given in [3]. Hence, in application of the CPL network to equalization, we can represent any nonlinear channel as a CPL function, and if we use a CPL network as an equalizer, then, the global system, cascade of nonlinear channel and the equalizer is still a CPL function since the class of CPL functions is closed.

To show the ability of CPL network to achieve blind equalization, we first introduce the following: The nonlinear channel $h(\cdot)$ maps the input sequence $x(n) \in \Omega$ to $y(n) = h(x(n), x(n-1), \dots, x(n-1-p))$ and the CPL

*Research supported in part by NSF grant NCR-9614236.

equalizer aims to recover the input sequence by constructing a mapping $h_{eq} : D \rightarrow \Omega$ where $D \subset R^k$ and $\Omega \subset R$. Assume that the global system, cascade of the nonlinear channel $h(\cdot)$ and the CPL equalizer $h_{eq}(\cdot)$ is denoted by $\mathcal{T}$, and is modeled by a CPL network which divides the input space into m disjoint regions, $R_1, R_2, \dots, R_m$, and in each region R_i , the CPL function given in (1) is equivalent to the following linear model:

$$M_i : \quad \tilde{x}(n) = \sum_{j=1}^k w_{ij} x_j(n) \quad (2)$$

where $x_j(n) \equiv x(n - j + 1)$ and $\tilde{x}(n)$ is the output of the equalizer.

We then make the following assumptions:

- (i) Input sequence $\{x(n)\}$ is an i.i.d. random process with distribution ν .
- (ii) The distribution ν is symmetric with finite variance.
- (iii) For each model M_i , there exists at least a subset $I \subset \Omega$ and a region R_i , $\tilde{I} = I_1 \times I_2 \times \dots \times I_k \subset R_j$ such that the mapping: $\tilde{x}(t) = \sum_{j=1}^k w_{ij} x_j(t)$ is a $\tilde{I}$ onto I mapping and $\rho = \int_{I+I'} d\nu_x = \int_{I_j+I_{j'}} d\nu_x = \rho_j, j = 1, 2, \dots, k$ where I' is the symmetric set of I .

We then show the following:

Theorem 2: Consider the global system $\mathcal{T} = \mathcal{F}(\mathcal{S})$. We assume that $\{x(n)\}$ is an i.i.d. process with distribution ν and assumptions (i), (ii), and (iii) are satisfied. If the distribution of $\{\tilde{x}(n)\}$ is still ν , then, the global system $\mathcal{T}$ is identity except for a possible delay and a sign factor.

Proof: Since $\{x(n)\}$ and $\{\tilde{x}(n)\}$ have the same distribution and ν is symmetric, then, $E\{x(n)\} = E\{\tilde{x}(n)\} = 0$ and

$$\begin{aligned} \int_{I+I'} x^2 d\nu_x &= \int_{\tilde{I}+\tilde{I}'} \tilde{x}^2 d\nu_{\tilde{x}} \\ &= \int_{I_1+I_{1'}} \dots \int_{I_k+I_{k'}} \tilde{x}^2 d\nu_{x_1} d\nu_{x_2} \dots d\nu_{x_k} \end{aligned}$$

where $I' \equiv -I$, $\tilde{I}' \equiv (-I_1) \times (-I_2) \times \dots \times (-I_k)$. Then, by (2), we can write

$$\int_{I+I'} x^2 d\nu_x = \sum_{j=1}^k \int_{I_1+I_{1'}} \dots \int_{I_k+I_{k'}} w_{ij}^2 x_j^2 d\nu_{x_1} d\nu_{x_2} \dots d\nu_{x_k}$$

which by using assumption (iii) gives,

$$\sigma_I^2 = \sum_{j=1}^k w_{ij}^2 \pi_j \sigma_j^2$$

where $\sigma_I^2 \equiv \int_{I+I'} x^2 d\nu_x$, $\sigma_j^2 \equiv \int_{I_j+I_{j'}} x_j^2 d\nu_{x_j}$, and $\pi_j \equiv \prod_{l \neq j} \rho_l$. Let $\gamma_j \equiv \pi_j \sigma_j^2 / \sigma_I^2$, we have

$$\sum_{j=1}^k w_{ij}^2 \gamma_j = 1. \quad (3)$$

Let f and f_j be the characterization functions [11] of $x(n)$ on $I+I'$ and $I_j+I_{j'}$ respectively. By the definition in (2), we have

$$f(\tau) = \prod_{j=1}^k f_j(w_{ij} \tau). \quad (4)$$

Let $g = |f|$ and $g_j = |f_j|$. From (4), $g(\tau) = \prod_{j=1}^k g_j(w_{ij} \tau)$. Setting $\psi(\tau) = -\ln g(\tau)/\tau^2$ and $\psi_j(\tau) = -\ln g_j(\tau)/\tau^2$, we have $\psi(\tau) = \sum_{j=1}^k w_{ij}^2 \psi_j(w_{ij} \tau)$ which can be rewritten as

$$\sum_{j=1}^k w_{ij}^2 [\gamma_j \psi(\tau) - \psi_j(w_{ij} \tau)] = 0 \quad (5)$$

by using (3). It follows from (5) that, for any τ , there exists at least one w_{ij} , such that $\gamma_j \psi(\tau) - \psi_j(w_{ij} \tau) \geq 0$. Then, we get

$$\gamma_j \geq \frac{\psi_j(w_{ij} \tau)}{\psi(\tau)} = \frac{\ln g_j(w_{ij} \tau)}{\ln g(\tau)} \frac{1}{w_{ij}^2}$$

Since $\rho = g(0)$, $\rho_j = g_j(0)$, and $\rho = \rho_j$ by assumption (iii), as τ goes to zero, we have

$$\gamma_j \geq \frac{1}{w_{ij}^2}. \quad (6)$$

From (3), we know that (6) holds if and only if $k = 1$. Thus, for each model M_i , there exists only one non-zero coefficient w_{ij_i} such that $w_{ij_i}^2 = 1$ and $\tilde{x}(n) = w_{ij_i} x_j(n - j_i + 1)$. Because of the continuity of the global CPL model, for any two models M_{i_1} and M_{i_2} that have a common boundary, we have $w_{ij_{i_1}} x_j(n - j_{i_1} + 1) = w_{ij_{i_2}} x_j(n - j_{i_2} + 1)$. Since n is varying, the above equality is true if and only if $w_{ij_{i_1}} = w_{ij_{i_2}}$ and $j_{i_1} = j_{i_2}$. Therefore, for all the models M_i , the time delay index j_i must be the same. This completes the proof of the theorem. $\square$

In the next section, the above conclusions are generalized to the RCPL network.

3. NONLINEAR BLIND EQUALIZATION BY RCPL NETWORK

We have introduced RCPL network in [9] and have shown that RCPL structure provides savings in computation and implementation, especially when required to model strong nonlinearities. Since RCPL network also employs feedback, it has a distinct dynamic behavior which is completely different from that attained by the use of finite duration impulse response feedforward structures. The RCPL function is defined as [9]:

Definition 2 (Recurrent Canonical Piecewise-Linear Function): A function $f: D_1 \times D_2 \times I \rightarrow Q$ with sample space $D_1 \subset R^N$, $D_2 \subset R^r$, index set I , and compact subset $Q \subset R^M$ is said to be a RCPL function if it can be expressed by the global representation:

$$f(\mathbf{x}(n), \mathbf{u}(n)) = \mathbf{a} + \mathbf{B}_1 \mathbf{x}(n) + \mathbf{B}_2 \mathbf{u}(n) \quad (7)$$

$$x_k(n) = a_k + \mathbf{b}_{1k}^T \mathbf{x}(n-1) + \mathbf{b}_{2k}^T f(\mathbf{x}(n-1), \mathbf{u}(n-1))$$

$$+ \mathbf{b}_{3k}^T \mathbf{u}(n) + \sum_{i=1}^{\tau} c_{ki} |\langle \alpha_{1ki}, \mathbf{x}(n-1) \rangle|$$

$$+ \langle \alpha_{2ki}, f(\mathbf{x}(n-1), \mathbf{u}(n-1)) \rangle + \langle \alpha_{3ki}, \mathbf{u}(n) \rangle + \beta_{ki} \quad (8)$$

where $\mathbf{x}, \mathbf{b}_{1k}, \alpha_{1ki} \in R^N$, $\mathbf{u}, \mathbf{b}_{3k}, \alpha_{3ki} \in R^r$, $\mathbf{a}, \mathbf{b}_{2k}, \alpha_{2ki} \in R^M$, $\mathbf{B}_1 \in R^{M \times N}$, $\mathbf{B}_2 \in R^{M \times r}$, $a_k, c_{ki}, \beta_{ki} \in R$, and $k = 1, 2, \dots, N$. x_k is the k th element in $\mathbf{x}$. We refer to the structure defined by (7) and (8) as the *recurrent canonical piecewise linear network*.

By comparing definitions 1 and 2, we can see that RCPL filter is actually a special case of the CPL filter. The RCPL

filter partitions the input signal space into finite disjoint regions and in each region, it can be represented by a FIR filter with infinite length. Therefore, the result presented in Theorem 1 also holds for the RCPL filter. Hence, any nonlinear channel can be represented as a RCPL function. Furthermore, if we use RCPL network as an equalizer, then, the global system which consists of the channel and equalizer is still a RCPL function. Let the global system $\mathcal{T}$ be a RCPL network and input variable $\{x(n)\}$ be an i.i.d. random variable with symmetric distribution ν . Then, we have the same conclusion given in Theorem 2 also for the RCPL filter. Thus, for blind equalization, we can update the weights of RCPL equalizer in such a way that the instantaneous distribution of the output $\hat{x}(n)$ of the equalizer converges to the input distribution ν . Several cost functions such as moment error cost function [8], Godard/Sato cost function [7], Vembu's convex cost function [10] and partial likelihood cost function [1], [2] can be used for distribution matching. However, moment error cost function is not a consistent function, especially in the nonminimum phase case [3]. We have shown that partial likelihood cost function can be successfully used for blind equalization of binary communication channels [9]. Godard/Sato cost function is not a convex cost function and the derived blind algorithm may only find a local minimum. However, if proper initial weights are chosen, the algorithm can reach the global minimum and the equalizer prediction error tends to zero. Thus, the resulting blind equalizer has larger stable margin. Vembu's convex cost function can help the algorithm to find the global minimum for linear channels and the equalizer to achieve the correct decision boundary however, it results in larger residual prediction error after convergence. Although we can not show that Vembu cost function is also a convex cost function for the nonlinear channel, it may help us to find an initial proper region in the weight space for the search, facilitated by the piecewise linear structure of the RCPL network. In the next section, we present application of the Godard and Vembu cost functions for equalization with the RCPL network and derive a blind algorithm based on both the Godard and Vembu cost functions, such that initially Vembu cost function is used in the learning process and when the absolute gradient of Godard error change becomes small, adaptation is switched to the Godard cost function.

4. IMPLEMENTATION OF BLIND ALGORITHM

We introduce decision feedback to the RCPL network which results in the final RCPL-DF equalizer structure shown in Figure 1. The decision feedback (DF) structure provides better performance especially when the channel characteristics involves nonminimum phase multipath components. This improvement in performance is also demonstrated by our simulation studies. The dynamics of RCPL-DF equalizer given in Figure 1 can be described by the following set of equations:

$$\begin{aligned} z_k(n) &= \omega_{k0}(n)\hat{x}(n-1) + \sum_{i=1}^M \omega_{ki}(n)\hat{x}_i(n-1) \\ &+ \sum_{i=1}^N \omega_{ki+M}(n)y(n-i+1) + \sum_{i=1}^Q \omega_{ki+M+N}(n)\bar{x}(n-i), \\ \hat{x}_k(n) &= f_k(z_k(n)), \quad k = 1, 2, \dots, M, \quad \bar{x}(n) = q(\hat{x}_k(n)), \\ \hat{x}(n) &= \sum_{i=1}^M \omega_{0i}(n)\hat{x}_i(n) + \sum_{i=1}^N \omega_{0i+M}(n)y(n-i+1) \end{aligned}$$

$$+ \sum_{i=1}^Q \omega_{0i+M+N}(n)\bar{x}(n-i) \quad (9)$$

where $y(n)$ is the observed channel output corresponding to the transmitted signal $x(n)$ which takes values from a finite set $\mathcal{S}$, $\hat{x}(n)$ is the output of a unit trained to approximate $x(n)$, $\bar{x}(n)$ is the signal after the decision function $q(\cdot)$, and $f_k(\cdot)$ is a piecewise-linear function. If we choose the functions $f_k(\cdot)$ as

$$f_k(z_k(n)) = |z_k(n) + 1| - |z_k(n) - 1| \quad k = 1, 2, \dots, M$$

for the Godard cost function $J_g(n)$ and Vembu's cost function $J_v(n)$:

$$J_g(n) = E\{|\hat{x}(n)|^p - R_p\}, R_p = \frac{E\{|x(n)|^{2p}\}}{E\{|x(n)|^p\}}, p = 1, 2, \dots$$

$$J_v(n) = E\{|\hat{x}(n)|^r\}, r = 2, 3, \dots$$

a learning algorithm can be obtained by steepest descent minimization of these cost functions:

$$\mathbf{w}_0(n+1) = \mathbf{w}_0(n) + \mu_1 e(n) \mathbf{x}(n) \quad (10)$$

$$\mathbf{w}_k(n+1) = \mathbf{w}_k(n) + \mu_2 e(n) \omega_{0k} v_k(n) \bar{\mathbf{x}}(n-1) \quad (11)$$

$$\text{where } e(t) = e_g(t) \quad \text{if } |e_g(t)| < \alpha; \quad \text{else } e(t) = e_v(t)$$

$$e_g(n) = \text{sgn}(\hat{x}(n))|\hat{x}(n)|^{p-1}(R_p - |\hat{x}(n)|^p)$$

$$e_v(t) = r \text{sgn}(\hat{x}(n))|\hat{x}(n)|^{r-1}$$

$$v_k(n) = \text{sgn}(z_k(n) + 1) - \text{sgn}(z_k(n) - 1)$$

where $\bar{\mathbf{x}}^T(n-1) = [\hat{x}(n-1), \mathbf{x}(n-1)]$, $\mathbf{x}^T(n) = [\hat{x}(n), \hat{x}_1(n), \dots, \hat{x}_M(n), y(n), \dots, y(n-N+1), \bar{x}(n-1), \dots, \bar{x}(n-Q)]$, $\mathbf{w}_j(n) = [\omega_{j0}, \omega_{j2}, \dots, \omega_{jN+M+Q}]$, $j = 0, 1, \dots, M$, and μ_1, μ_2 are the learning rates.

As an example, consider the nonlinear communication channel $g(n) = g_l(n) + 0.1g_l(n)^2$ where the nonminimum phase multipath component is given by $g_l(n) = 0.9\delta(n) + \delta(n-1)$. The input $x(n)$ takes values from the binary set $\mathcal{S} = \{-1, 1\}$ and has a symmetric distribution. We choose $M = 3$, $N = 5$, $Q = 3$, $p = 2$, $r = 4$. Figure 2 shows the probability of decision error of RCPL-DF equalizer for Vembu, Godard, and the equalizer given by (10)–(11), the combined Godard and Vembu cost functions. Here, input signal to noise ratio (SNR) is 25 dB, $\alpha = 15$, $\mu_1 = 0.001$, $\mu_2 = 0.01$ for the Godard cost function and $\mu_1 = 0.00006/r$, $\mu_2 = 0.005/r$ for the Vembu cost function. The mean square error (MSE) and bit error rate (BER) curves for the same three equalizers are shown in Figure 3. We can see that the blind algorithm based on the combined cost function is faster than the ones based on the Godard or Vembu cost function alone. However, after convergence, all three exhibit comparable BER performance. It is also important to note that the algorithm convergence is much faster when the multipath component $g_l(n)$ is chosen as minimum phase. Figure 4 shows the BER curve after 8000 iterations for the same channel for CMA, linear-DF, and RCPL-DF equalizers using the Godard cost function. The length of the linear filters is chosen as 15 and the length of the DF section for the RCPL-DF and linear-DF equalizers as 3. As observed in the figure, the linear CMA equalizer exhibits very poor performance for the given nonlinear channel while both the RCPL-DF and linear-DF achieve quite satisfactory equalization with RCPL-DF outperforming the linear-DF equalizer. All the simulation results presented are averaged over 25 independent runs. It is also worth noting that the linear-DF equalizer can be also regarded as a special case of the RCPL equalizer and the results we present in this paper can be extended to this equalizer structure.

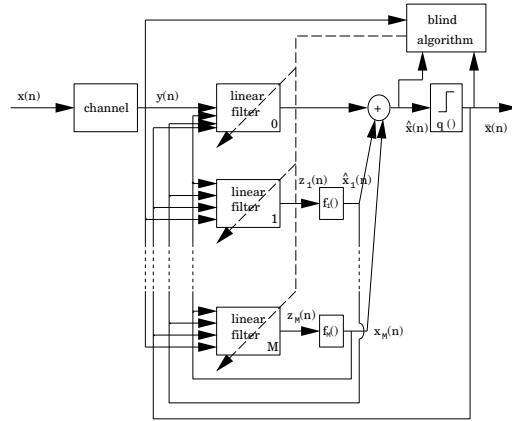


Figure 1. Decision feedback-RCPL equalizer structure

REFERENCES

- [1] T. Adalı, X. Liu, and M. K. Sönmez, "Conditional distribution learning with neural networks and its application to channel equalization," to appear *IEEE Trans. Signal Processing*.
- [2] T. Adalı, M. K. Sönmez, and X. Liu, "Partial likelihood estimation for real-time signal processing," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)*, Atlanta, GA, May 1996, vol. 6, pp. 3562-3565.
- [3] T. Adalı and X. Liu, "Canonical piecewise linear network for nonlinear filtering and its application to blind equalization," submitted to *Signal Processing*.
- [4] S. Benedetto, E. Biglieri, and V. Castellani, "Digital Transmission Theory," Prentice Hall Int. Englewood Cliffs, New Jersey, 1987.
- [5] L. O. Chua and A. Deng, "Canonical piecewise-linear networks," *IEEE Trans. Circuits Syst.*, vol. 35, No. 1, pp. 101-111, Jan. 1988.
- [6] D. Donoho, "On minimum entropy deconvolution," in *Applied time series analysis II*, ed. D. Findley, Academic Press, New York, pp. 556-608.
- [7] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. COM-28, pp. pp. 1867-1875, Nov. 1980.
- [8] G. Kechriotis, E. Zervas, and E. S. Manolakos, "Using recurrent neural networks for adaptive communication channel equalization," *IEEE Trans. Neural Networks*, vol. 5, no. 2, pp. 267-278, March 1994.
- [9] X. Liu and T. Adalı, "Channel equalization using partial likelihood estimation and recurrent canonical piecewise-linear network," in *Proc. EUSIPCO-96, VIII European Signal Processing Conf.*, Trieste, Italy, Sept. 1996, pp. 383-386.
- [10] S. Vembu, S. Verdu, R. A. Kennedy, and W. Sethares, "Convex cost functions in blind equalization," *IEEE Trans. Signal Processing*, vol. 42, No. 8, pp. 1952-1959, Aug. 1994.
- [11] A. M. Kagan, Y. V. Linnik, and C. R. Rao, "Characterization Problems in Mathematical Statistics," John Wiley & Sons, Inc., 1973.

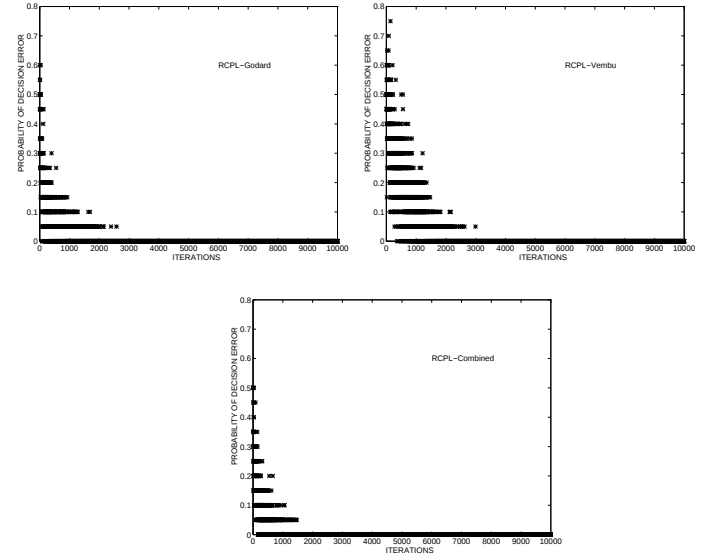


Figure 2. Convergence curves for the RCPL equalizer for the three cost functions

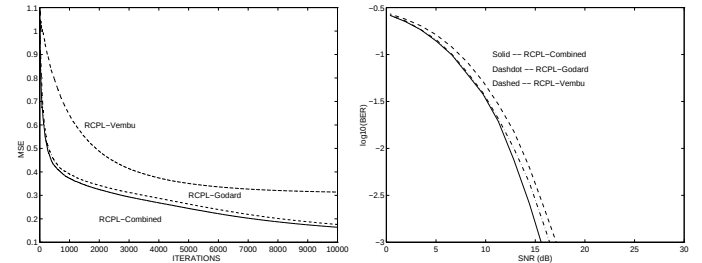


Figure 3. MSE and BER curves for the RCPL equalizer for the three cost functions

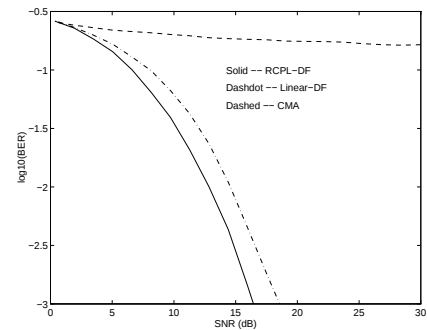


Figure 4. BER curves for linear (CMA), linear-DF, and RCPL-DF equalizers using Godard cost function