

Convergence Analysis of a Prediction based Adaptive Equalizer for IIR Channels

Miloje S. Radenkovic, and Tamal Bose

Abstract– This paper presents the convergence analysis of a prediction based blind equalizer for IIR channels. Predictor parameters are estimated by using the recursive least squares algorithm. It is shown that the prediction error converges almost surely (a.s.) toward a scalar multiple of the unknown input symbol sequence. It is also proved that the convergence rate of the parameter estimation error is of the same order as that in the iterated logarithm law.

Keywords– Adaptive blind equalizer, Recursive least squares, Adaptive Filtering, Convergence analysis.

I. INTRODUCTION

THE earliest blind channel equalization methods were based on single-input single-output models, sampled at the symbol rate. Some of them such as the constant modulus algorithms (CMA) involve nonlinear optimization and higher order statistics of the channel output [1], [2]. An extensive list of references of CMA methods is given in [3]. Since the appearance of [4], a large number of blind equalization results are based on using only second order statistics of the received signals, (see for example [5], [6] and [7]). The basic idea in [4] is to oversample the received signal with respect to the baud rate, or to use multiple antennas thereby giving single-input multiple-output (SIMO) channel model. For a comprehensive list of important contributions in this area until 1998, we refer to [3] and [7].

In this paper we present an RLS based blind adaptive equalizer, where the parameter estimates are updated when each single signal is received. Following the idea presented in [6], [8], [9], an equalizer is developed based on one-step ahead prediction of the received signal. It is shown that IIR channels can be equalized with the FIR type predictors. It is proved that almost surely (a.s.)

This work was supported in part by the State of Utah Centers of Excellence Program and NASA Grant NAG5-10716.

Miloje S. Radenkovic is with the Electrical Engineering Department, University of Colorado, Denver, Colorado 80127.

Tamal Bose is with the Center for High-speed Information Processing (CHIP), Department of Electrical and Computer Engineering, Utah State University, Logan, Utah 84322.

(with probability one) the prediction error converges to a scalar multiple of the input symbol sequence. Rigorous analysis reveals that the convergence rate of the parameter estimation error is the same as that in the iterated logarithm law.

II. PROBLEM STATEMENT

The standard model of a fractionally spaced receiver is the single-input multiple-output system. For simplicity of our presentation, we consider a single-input two-output system model. In this case the receiver performs the two measurements,

$$\begin{aligned} x_1(i) &= q^{-d} \frac{B_1(q^{-1})}{A_1(q^{-1})} w(i), \\ x_2(i) &= q^{-d} \frac{B_2(q^{-1})}{A_2(q^{-1})} w(i) \end{aligned}$$

for each transmitted symbol $w(i)$, $i \geq 0$. Here q^{-1} is the unit delay operator, integer d is the delay between input $w(i)$ and outputs $x_k(i)$, $k = 1, 2$, and $B_1(q^{-1})/A_1(q^{-1})$ and $B_2(q^{-1})/A_2(q^{-1})$ are stable IIR transfer operators, and $B_i(q^{-1})$ and $A_i(q^{-1})$, $i = 1, 2$ are polynomials in q^{-1} . An equivalent representation of this process is given in Fig. 1, where

$$\begin{aligned} x_1(i) &= q^{-d} \frac{B(q^{-1})}{A(q^{-1})} w(i), \\ x_2(i) &= q^{-d} \frac{C(q^{-1})}{A(q^{-1})} w(i), \end{aligned} \quad (1)$$

with

$$\begin{aligned} A(q^{-1}) &= 1 + a_1 q^{-1} + \dots + a_L q^{-L}, \\ B(q^{-1}) &= b_0 + b_1 q^{-1} + \dots + b_L q^{-L}, \\ C(q^{-1}) &= c_0 + c_1 q^{-1} + \dots + c_L q^{-L}, \end{aligned}$$

where L is the channel order. Clearly, if $A(q^{-1}) = A_1(q^{-1})A_2(q^{-1})$, then $B(q^{-1}) = B_1(q^{-1})A_2(q^{-1})$, and $C(q^{-1}) = B_2(q^{-1})A_1(q^{-1})$. In general $w(i)$, $x_1(i)$, $x_2(i)$ and the coefficients of polynomials in (1) can be complex quantities.

In this paper it is assumed that $\frac{1}{A(q^{-1})}$ is a stable operator. We introduce the following assumptions regarding the channel model of Fig. 1.

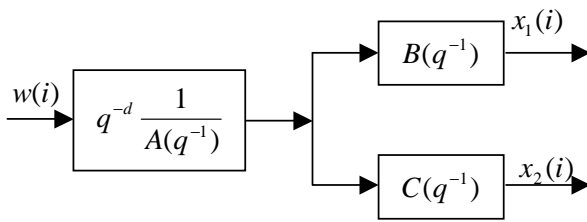


Fig. 1 Signal Model

Assumption A1: $B(q^{-1})$ and $C(q^{-1})$ are coprime polynomials.

Assumption A2: Let $w(i) = w_R(i) + jw_I(i)$, $j = \sqrt{-1}$, with $\{w_R(i)\}$ and $\{w_I(i)\}$ being real sequences. Let $\{w(i)\}$ be a martingale difference sequence, i.e.,

$$E \left[\begin{pmatrix} w_R(i+1) \\ w_I(i+1) \end{pmatrix} \middle| F_i \right] = 0, \quad (a.s.) \quad (2)$$

where, $F_i \triangleq \{w_R(0), \dots, w_R(i), w_I(0), \dots, w_I(i)\}$,

$$E(w(i+1)w(i+1)^* | F_i) = \sigma_w^2, \quad (a.s.) \quad (3)$$

where $(.)^*$ denotes complex conjugate and

$$|w(i)| \leq k_w < \infty. \quad (4)$$

Note that (2) implies

$$\begin{aligned} & E(w(i)w(i+l)^*) \\ &= E[E(w(i)w(i+l)^* | F_i)] \\ &= E[w(i)w(i+l)^* | F_i] = 0, \quad l \geq 1. \end{aligned} \quad (5)$$

That is, samples of $w(i)$ at two different time instants are uncorrelated. Observe that A2 does not imply that the samples of $w(i)$ are independent random variables.

The prediction based equalizer is described in Fig. 2, where $y(i+1)$ is one step-ahead prediction of $x_1(i+1)$, based on the observed samples $x_1(k)$ and $x_2(k)$, $k \leq i$. Hence

$$y(i+1) = R(q^{-1})x_1(i) + S(q^{-1})x_2(i), \quad (6)$$

where the filter operators are defined as

$$\begin{aligned} R(q^{-1}) &= r_0 + r_1q^{-1} + \dots + r_{N_1}q^{-N_1}, \\ S(q^{-1}) &= s_0 + s_1q^{-1} + \dots + s_{N_2}q^{-N_2}, \end{aligned} \quad (7)$$

$$N_1 \geq L-1, \quad N_2 \geq L-1$$

where L is the channel order. We consider the predictor to be optimal in the mean-square sense. The order of each of the polynomials $R(q^{-1})$ and $S(q^{-1})$ is $L-1$, the reason for which will be given shortly. These polynomials are computed by minimizing the following cost function:

$$J = E(|x_1(i+1) - y(i+1)|^2). \quad (8)$$

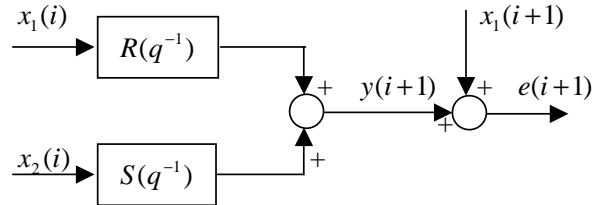


Fig. 2 Predictor based equalizer

Note that instead of $x_1(i)$ we can use $x_2(i)$ just as well as the reference signal and derive the predictor by minimizing $E(|x_2(i+1) - y(i+1)|^2)$. Next we explain how to calculate $R(q^{-1})$ and $S(q^{-1})$ so that J in (8) is minimal.

Note that $x_1(i+1)$ can be written in the form,

$$\begin{aligned} x_1(i+1) &= b_0w(i+1-d) \\ &+ \frac{q[B(q^{-1}) - A(q^{-1})b_0]}{A(q^{-1})}w(i-d), \end{aligned} \quad (9)$$

where q is forward shifting operator, and b_0 is the leading coefficient of $B(q^{-1})$. Since from (8) and Fig. 2,

$$\begin{aligned} y(i+1) &= P(q^{-1})\frac{B(q^{-1})}{A(q^{-1})}w(i-d) \\ &+ S(q^{-1})\frac{C(q^{-1})}{A(q^{-1})}w(i-d), \end{aligned} \quad (10)$$

equation (9) yields,

$$x_1(i+1) - y(i+1) = x(i-d) + b_0w(i+1-d), \quad (11)$$

$$\begin{aligned} x(i-d) &= \left[\frac{B(q^{-1})R(q^{-1}) + C(q^{-1})S(q^{-1})}{A(q^{-1})} - \right. \\ &\quad \left. \frac{q(B(q^{-1}) - A(q^{-1})b_0)}{A(q^{-1})} \right]w(i-d), \end{aligned} \quad (12)$$

where for the sake of notational simplicity, q^{-1} is out shifted in the above polynomials. Using the fact that by (5) samples of $\{w(i)\}$ are uncorrelated, from (11) it follows that J is minimal if and only if $x(i-d) = 0$. Then

$$x_1(i+1) - y(i+1) = b_0w(i+1-d), \quad (13)$$

and $\min J = |b_0|^2\sigma_w^2$. Relation (12) implies that $x(i) = 0$ if

$$\begin{aligned} & B(q^{-1})R(q^{-1}) + C(q^{-1})S(q^{-1}) \\ &= q(B(q^{-1}) - A(q^{-1})b_0), \end{aligned} \quad (14)$$

which is the design equation for calculating polynomials $R(q^{-1})$ and $S(q^{-1})$ of the optimal predictor (8). The degrees of $R(q^{-1})$ and $S(q^{-1})$ must be $L - 1$ as given in (7) so that the above equation has a unique solution. In practice, $B(q^{-1})$ and $C(q^{-1})$ are unknown. Hence we cannot use (14) to calculate $R(q^{-1})$ and $S(q^{-1})$. We now propose a recursive algorithm for directly estimating the unknown parameters in (8). Let

$$\theta_0^H = [r_0, r_1, \dots, r_{N_1}, s_0, s_1, \dots, s_{N_2}], \quad (15)$$

and

$$\phi(i)^T = [x_1(i), x_1(i-1), \dots, x_1(i-N_1), x_2(i), x_2(i-1), \dots, x_2(i-N_2)], \quad (16)$$

where $(\cdot)^H$ stands for conjugate transpose. Then from (6) we have,

$$y(i+1) = \theta_0^H \phi(i). \quad (17)$$

Instead of (17) we use following adaptive predictor,

$$\hat{y}(i+1) = \hat{\theta}(i)^H \phi(i), \quad (18)$$

where $\hat{\theta}(i)$ is an estimate of unknown θ_0 , and it can be generated by the following recursive least square (RLS) algorithm [10]:

$$\hat{\theta}(i+1) = \hat{\theta}(i) + p(i)\phi(i)\epsilon(i+1)^*, \quad (19)$$

$$\epsilon(i+1) = x_1(i+1) - \hat{\theta}(i)^H \phi(i), \quad (20)$$

$$p(i) = p(i-1) - \frac{p(i-1)\phi(i)\phi(i)^H p(i-1)}{1 + \phi(i)^H p(i-1)\phi(i)},$$

$$p(0) = p_0 I, p_0 > 0, \quad (21)$$

where I is the identity matrix.

III. CONVERGENCE OF THE ADAPTIVE EQUALIZER

We now show consistency of the parameter estimates, and prove that the prediction error converges almost surely (a.s.) towards the scalar version of the input symbol sequence. In the following, $f(i) = \mathbf{O}(g(i))$ means $\lim_{i \rightarrow \infty} |f(i)/g(i)| < \infty$.

Theorem 1 *Let Assumptions A1 and A2 hold and one of the following is valid: $N_1 = L - 1$ or $N_2 = L - 1$. Then,*

$$\|\hat{\theta}(i) - \theta_0\| = \mathbf{O}\left(\frac{\log \log i}{i}\right)^{\frac{1}{2}} \quad (a.s.) \text{ as } n \rightarrow \infty, \quad (22)$$

and

$$\lim_{i \rightarrow \infty} (\epsilon(i+1) - b_0 w(i+1-d)) = 0. \quad (a.s.) \quad (23)$$

Proof: Note that by the Matrix Inversion Lemma [10], (21) gives

$$p(i)^{-1} = p(i-1)^{-1} + \phi(i)\phi(i)^H, \quad (24)$$

or

$$p(i)^{-1} = p(0)^{-1} + \sum_{k=1}^i \phi(k)\phi(k)^H. \quad (25)$$

We first prove that for sufficiently large i , $\frac{p(i)^{-1}}{i}$ is a positive definite matrix. Since from (1), $x_1(i) = \frac{B}{A}w(i-d)$ and $x_2(i) = \frac{C}{A}w(i-d)$, signal vector $\phi(i)$ given by (16), can be written as follows:

$$\phi(i)^T = \frac{1}{A(q^{-1})} [B(q^{-1}), q^{-1}B(q^{-1}), \dots, q^{-N_1}B(q^{-1}), C(q^{-1}), q^{-1}C(q^{-1}), \dots, q^{-N_2}C(q^{-1})] w(i-d) \quad (26)$$

Using the fact that by Assumption A2, process $w(i)$ is ergodic in the second moment, application of Parseval's theorem on (25) gives

$$T = \lim_{i \rightarrow \infty} \frac{p(i)^{-1}}{i} = \frac{1}{2\pi} \int_0^{2\pi} F(e^{-j\omega}) F(e^{j\omega})^H \sigma_w^2 d\omega \quad (a.s.)$$

where

$$F(z^{-1})^T = \frac{z^{-d}}{A(z^{-1})} [B(z^{-1}), z^{-1}B(z^{-1}), \dots, z^{-N_1}B(z^{-1}), C(z^{-1}), z^{-1}C(z^{-1}), \dots, z^{-N_2}C(z^{-1})], z = e^{j\omega}.$$

Obviously T is positive definite if it is not possible to find any nonzero vector λ satisfying $\lambda^T T \lambda = 0$. The last equality holds if and only if

$$\lambda^T F(z^{-1}) = 0, \quad (27)$$

for all $\omega \in [0, 2\pi]$. Let $\lambda_1(z^{-1})$ and $\lambda_2(z^{-1})$ be polynomials in $z = e^{j\omega}$ of degree N_1 and N_2 respectively, formed from the corresponding components of λ . Then (27) becomes

$$\lambda^T F(z^{-1}) = \lambda_1(z^{-1})B(z^{-1}) + \lambda_2(z^{-1})C(z^{-1}) = 0. \quad (28)$$

Since $B(z^{-1})$ and $C(z^{-1})$ are coprime, and either $N_1 = L-1 < \deg B(z^{-1}) = \deg C(z^{-1}) = L$ or $N_2 = L-1 < L$, (28) holds if $\lambda_1(z^{-1}) = 0$ and $\lambda_2(z^{-1}) = 0$ implying $\lambda = 0$, and therefore T is a positive definite matrix. Next we prove statement (22). Since from (13) and (17) $x_1(i+1) = y(i+1) + b_0 w(i+1-d) = \theta_0^H \phi(i) + b_0 w(i+1-d)$, we have

$$\begin{aligned} \epsilon(i+1) &= x_1(i+1) - \hat{\theta}(i)^H \phi(i) \\ &= -\hat{\theta}(i)^H \phi(i) + b_0 w(i+1-d), \end{aligned} \quad (29)$$

where $\tilde{\theta}(i)$ is parameter estimation error given by $\tilde{\theta}(i) = \hat{\theta}(i) - \theta_0$. By using (29) in (19) it follows that

$$\tilde{\theta}(i+1) = \tilde{\theta}(i) + p(i)\phi(i)[- \phi(i)^H \tilde{\theta}(i) + b_0^* w(i+1-d)^*], \quad (30)$$

or,

$$p(i)^{-1} \tilde{\theta}(i+1) = p(i)^{-1} \tilde{\theta}(i) - \phi(i)\phi(i)^H \tilde{\theta}(i) + b_0^* \phi(i) w(i+1-d)^*. \quad (31)$$

Substituting (24) into the first term on the right hand side of (31) yields

$$p(i)^{-1} \tilde{\theta}(i+1) = p(i-1)^{-1} \tilde{\theta}(i) + b_0^* \phi(i) w(i+1-d)^*, \quad (32)$$

from where we obtain

$$p(i)^{-1} \tilde{\theta}(i+1) = p(0)^{-1} \tilde{\theta}(1) + b_0^* \sum_{k=1}^i \phi(k) w(k+1-d)^*. \quad (33)$$

Consider the following martingale transform,

$$S(i+1-d) \triangleq \sum_{k=1}^n \phi(k) w(k+1-d)^*; \quad i \geq d-1.$$

Then

$$E(S(i+1-d) | F_{i-d}) = S(i-d), \quad (a.s.) \quad (34)$$

where we have used the fact that $\phi(i)$ depends only on past samples of $w(k)$, $k \leq i-d$. On the other hand since $\{\phi(i)\}$ is a bounded sequence, we have $E(|\phi(i)w(i+1-d)|) < \infty$. Hence (34) implies that $S(i+1-d)$ is a martingale. By virtue of the iterated logarithm law for martingales [11] we can conclude that

$$\|S(i+1-d)\| = \mathbf{O} \left[(i \log \log i)^{\frac{1}{2}} \right]. \quad (a.s.) \quad (35)$$

Using the fact that for sufficiently large i , $\frac{p(i)^{-1}}{i}$ is a positive definite matrix, statement (22) follows from (33) and (35). Furthermore, (22) implies $\lim_{i \rightarrow \infty} (\tilde{\theta}(i)^H \phi(i)) = 0$ (a.s.). Then statement (23) follows from (29) and the proof is complete. ■

Note that the convergence rate in (22) is the best possible for the parameter error generated by the least squares based algorithm, since it is the same as that in the laws of the iterated logarithm.

IV. SIMULATION EXAMPLE

In this experiment we use a symbol sequence generated from a 16-QAM constellation. The corresponding symbol levels along both axis are $-1.5, -0.5, 0.5$ and

1.5 . We consider the following channel model $A(q^{-1}) = 1 - 0.8q^{-1} + 0.16q^{-2}$, $B(q^{-1}) = (1.5 - 1.5i) + (-2.6 - 0.9i)q^{-1} + (3.44 + 1.2i)q^{-2} + (-2.32 + 1.7i)q^{-3}$, $C(q^{-1}) = (1 - 0.7i) + (-1.7 + 2i)q^{-1} + (0.96 - 1.4i)q^{-2} + 2.18q^{-3}$. In (16) we take $N_1 = 3$ and $N_2 = 4$. Fig. 3 shows the received symbols $x_1(i)$, while Fig. 4 presents the equalized symbols eye diagram. The amount of rotation and magnification in the eye diagram is a function of $b_0 = 1.5 - 1.5i$, i.e., the angle of rotation is -45° , while the magnification is $|b_0| = 2.1213$. According to (23),

$$ms(i) = \frac{1}{i} \sum_{k=1}^i |\epsilon(k) - b_0 w(k)|^2 \xrightarrow{i \rightarrow \infty} 0, \quad (a.s.) \quad (36)$$

Fig. 5 illustrates this fact for this particular example.

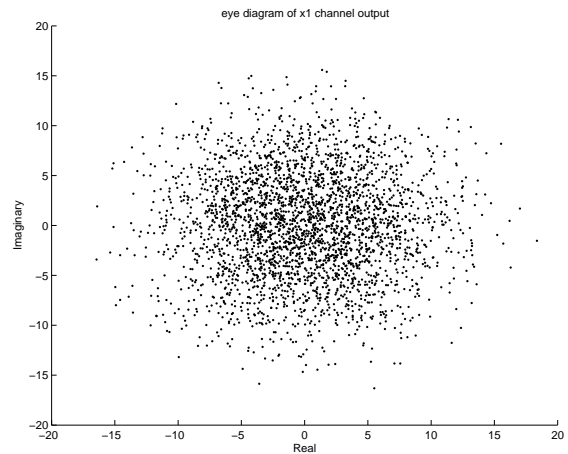


Fig. 3 Eye diagram of received signal x_1

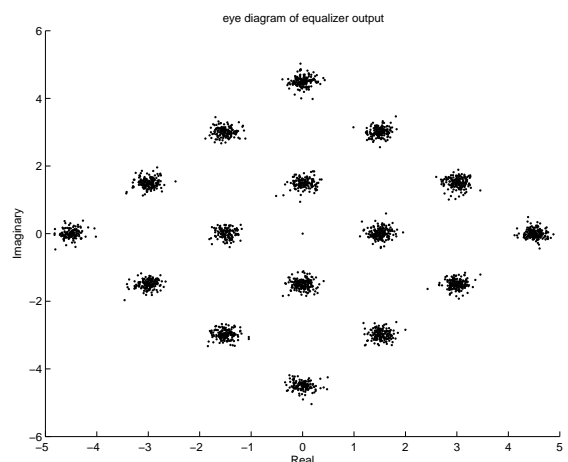


Fig. 4 Eye diagram of equalizer output

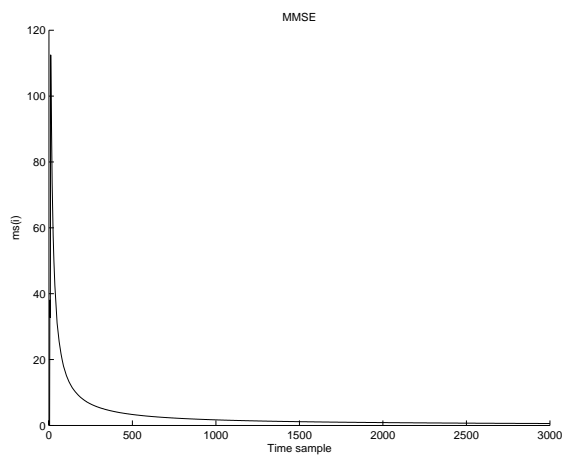


Fig. 5 Mean square error

V. CONCLUSION

This paper presents a rigorous convergence analysis of the prediction based RLS adaptive equalizer. It is proved that the parameter estimates converge (a.s.) with the same rate as that given by the iterated logarithm law. Also, the prediction error converges (a.s.) to a scalar version of the input symbol sequence. Currently, efforts are under way to extend the above results to the case when receiver noise is present.

REFERENCES

- [1] D. N. Godard, "Self recovering equalization and carrier tracking in the two dimensional data communication systems," *IEEE Trans. Commun.*, vol. COMM-28, pp. 1867-1875, November 1980.
- [2] O. Shalvi and E. Weinstein, "New criterion for blind deconvolution of minimum phase system (channels)," *IEEE Trans. Inform. Theory*, vol. 36, pp. 312-321, March 1990.
- [3] C. R. Johnson, P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown and R. A. Casas, "Blind equalization using the constant modulus criterion: a review," *Proc. of the IEEE*, vol. 86, no. 10, pp. 1927-1949, Oct. 1998.
- [4] L. Toug, G. Xu and T. Kailath, "Blind identification and equalization based on a second order statistics: a time domain approach," *IEEE Trans. Inform. Theory*, vol. 40, pp.340-349, March 1994.
- [5] E. W. Bai and M. Fu, "Blind system identification and channel equalization of IIR system without statistical information," *IEEE Trans. Signal Processing*, vol. 47, pp. 1910-1920, July 1999.
- [6] D. Gesbert and P. Duhamel, "Unbiased blind adaptive channel identification and equalization," *IEEE Trans. Signal Processing*, vol. 48, pp.148-158, Jan. 2000.
- [7] L. Tong and S. Perrean, "Multichannel blind identification: from subspace to maximum likelihood methods," *Proc. IEEE*, vol. 86, no. 10, pp.1951-1968, 1998.
- [8] K. Abed-Meriau, E. Mouliues and P. Loubaton "Prediction error method for second-order blind identification," *IEEE Trans. Signal Processing*, vol. 45, pp. 694-705, 1997.
- [9] J. Tugnait, "Multistage linear prediction based blind equalization of FIR/IIR single-input multiple-output channels with common zeros," *IEEE Trans. Signal Processing*, vol. 47, pp. 1689-1700, 1999.
- [10] S. Haykin, "Adaptive Filter Theory," Prentice Hall, 2002.
- [11] Y. S. Chow and H. Teicher, "Probability theory, independence, interchangibility, martingales," *Springer-Verlag*, New York, 1978.