Forty-Fourth Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
Sept 27-29, 2006

Tracking Performance of an LMS-Linear Equalizer for Fading Channels

Veeraruna Kavitha and Vinod Sharma

Abstract— **We consider a time varying wireless fading channel, equalized by an LMS linear equalizer. We study how well this equalizer tracks the optimal Wiener equalizer. We model the channel by an Auto-regressive (AR) process. Then the LMS equalizer and the AR process are jointly approximated by the solution of a system of ODEs (ordinary differential equations). Using these ODEs, the error between the LMS equalizer and the instantaneous Wiener filter is shown to decay exponentially/polynomially to zero unless the channel is marginally stable in which case the convergence may not hold. Using the same ODEs, we also show that the corresponding Mean Square Error (MSE) converges towards minimum MSE (MMSE) at the same rate for a stable channel. We further show that the difference between the MSE and the MMSE does not explode with time even when the channel is unstable. Finally we obtain an optimum step size for the linear equalizer in terms of the AR parameters, whenever the error decay is exponential.**

Key words: Fading channels, LMS Linear equalizer, tracking performance, ODE approximation.

I. INTRODUCTION

A channel equalizer is an important component of a communication system and is used to mitigate the ISI (inter symbol interference) introduced by the channel. The equalizer depends upon the channel characteristics. In a wireless channel, due to multipath fading, the channel characteristics change with time. Thus it may be necessary for the channel equalizer to track the time varying channel in order to provide reasonable performance.

Least Mean Square linear equalizer (LMS-LE) is a simple equalizer and is extensively used ([1], [3]). For a fixed channel its convergence to the Wiener filter has been studied in [1], [10] (see also the references therein). For a time varying channel, tracking behavior (how well LMS-LE tracks the instantaneous Wiener filter) of the algorithm is also important ([3]). However, although tracking of a channel estimator has been extensively studied (see, e.g., convergence analysis in [1], [6], [9], [14], and variable step size algorithms in [2], [7], [13] and references therein), we are not aware of theoretical studies on the tracking behavior of the LMS-LE (although various studies have compared their behavior with other schemes via simulations, approximations and upper bounds on probability of error and found that the tracking behavior of an LMS-LE may not be comparable to some others, see, e.g., [3], [5], [11]). Since LMS-LE is an important equalizer, it is certainly useful to obtain its tracking behavior theoretically. For example our study will show a dependence of its tracking behavior on certain system parameters which might not have been known before.

To study the tracking behavior theoretically, one needs to have a theoretical model of the fading channel. Auto Regressive (AR) processes have been shown to model such channels quite satisfactorily ([8], [12]). Infact, it is sufficient to model the fading channel by an AR(2) process ([5], [8], [12]). Therefore in this paper we limit ourselves to AR(2) processes. We will show that the trajectory of an AR(2) process can be approximated by a set of ODEs. Using these ODEs we show that the AR process can be approximated by an exponential/polynomial or a cosine waveform. We will also approximate the trajectory of the corresponding LMS-LE by another ODE. Next we will study the error process of the difference between the LMS-LE and the corresponding Wiener filter using the ODEs obtained and show that the error decays exponentially/polynomially with time unless the channel is marginally stable in which case it may not converge. We also show that the difference between MSE and the MMSE (MSE corresponding to the Wiener filter) decays exponentially at the same rate for a stable channel and does not explode with time for an unstable channel . Later on, for the cases, where the decay is exponential (this includes all stable and a class of unstable AR(2) processes), we obtain an optimum step size for the linear equalizer.

The paper is organized as follows. In Section II we explain our model. Section III provides the ODE approximation for the processes of interest. Section IV uses the ODE approximation to provide exponential/polynomial decay of the difference between the LMS-LE and the Wiener filter. We also obtain an optimum step size in this section. Section V provides examples to demonstrate the ODE approximations. Section VI concludes the paper. The appendices contain some details on the proofs.

II. SYSTEM MODEL AND NOTATIONS

We consider a system consisting of a wireless channel followed by an adaptive linear equalizer. At time k, the input to the channel is s_k, the channel impulse response is Z_k (assumed a time varying finite impulse response linear filter with length, L), and the channel output is y_k where,

$$y_k = \sum_{i=0}^{L-1} Z_{k,i} s_{k-i} + n_k.$$

Here $Z_{k,i}$ is the i^{th} component of Z_k and n_k is a zero mean AWGN with variance σ_n^2. We also assume s_k to be an IID

Veeraruna Kavitha is a PhD student in ECE Dept, IISc, Bangalore, India kavitha@pal.ece.iisc.ernet.in

Vinod Sharma is a Professor at ECE Dept, IISc, Bangalore, India vinod@ece.iisc.ernet.in

This research is partially supported by DRDO-IISc program on advanced research in Mathematical Engineering.

sequence from a finite alphabet with zero mean which is independent of n_k and Z_k.

The adaptive equalizer at time k, is an FIR linear filter θ_k of length M. The linear equalizer update equation is ,

$$\theta_{k+1} = \theta_k - \mu Y_k(\theta_k^T Y_k - s_k). \quad (1)$$

where, the length M channel output vector Y_k can be written as $Y_k = \pi_k S_k + N_k$, where S_k, N_k are the appropriate length input and noise vectors respectively. We assume $E\left[S_k S_k^T\right] = I$. The convolutional matrix π_k depends upon channel values $Z_k, \cdots, Z_{k-M+1}$ and is given by

$$\begin{bmatrix} Z_{k,1} & Z_{k,2} & \cdots & Z_{k,L} & 0 & \cdots & 0 \\ 0 & Z_{k-1,1} & \cdots & Z_{k-1,L-1} & & \cdots & 0 \\ 0 & \vdots & & & & & \\ 0 & 0 & \cdots & Z_{k-M+1,1} & & \cdots & Z_{k-M+1,L} \end{bmatrix}.$$

We model the channel update process by an AR(2) process,

$$Z_{k+1} = d_1 Z_k + d_2 Z_{k-1} + \mu W_k \quad (2)$$

where W_k is an IID white noise error sequence, independent of $\{s_k\}$ and $\{n_k\}$ processes. We assume $E|W_k|^4 < \infty$. An AR(2) process can approximate a wireless channel quite realistically ([12], [5]). The class of AR(2) processes include the Random Walk model, the Filtered Random Walk model (when $d_1 + d_2 = 1$ and $|d_2| \leq 1$) and the Autoregressive Second Order model ($d_1 = 2\rho cos w_0$, $d_2 = -\rho^2$ with ρ and w_0 representing the degree of damping and the dominating frequency respectively) ([8]). The analysis that follows considers an AR(2) process. The general ideas of our approach can be extended to an AR(n) process, $n > 2$ (see [4] for comments).

III. ODE Approximation

Consider the AR(2) process (2). Define $Z_t = Z_{[\frac{t}{\mu^n}]}$, $t \geq 0$, where $[x]$ represents the integer part of x. Here $n = 1$ for $d_2 \in (-1, 1]$ and $n = 1/2$ if $d_2 = -1$. In Appendix A we will show that for μ small, for any $T < \infty$, $\{Z_t, 0 \leq t \leq T\}$ can be approximated by the solution of the ODE,

$$\begin{aligned} \dot{Z}(t) &= \tfrac{1}{1+d_2}\left[E(W_1) + \eta Z(t)\right] \quad if \quad d_2 \in (-1,1], \\ \tfrac{d^2 Z(t)}{dt^2} &= E(W_1) + \eta Z(t) \quad if \quad d_2 = -1, \end{aligned} \quad (3)$$

where $\eta \stackrel{\triangle}{=} \frac{d_1+d_2-1}{\mu}$, when $Z_0 = Z_{-1} = Z(0)$ (and $\dot{Z}(0) = 0$ for $d_2 = -1$). The solution of the above ODEs is given by,

$$Z(t) := \begin{cases} C_1 e^{\frac{\eta}{1+d_2}t} & -\frac{E(W_1)}{\eta} \quad \eta \neq 0, d_2 \in (-1,1], \\ \frac{E(W_1)}{1+d_2} t & +C_1 \quad \eta = 0, d_2 \in (-1,1], \\ C_1 cosh(\sqrt{\eta}\ t) & -\frac{E(W_1)}{\eta} \quad \eta > 0, d_2 = -1, \\ C_1 cos(\sqrt{|\eta|}\ t) & +\frac{E(W_1)}{\eta} \quad \eta < 0, d_2 = -1, \\ \frac{E(W_1)}{2}\ t^2 & +C_1 \quad \eta = 0, d_2 = -1. \end{cases} \quad (4)$$

Now we consider the linear equalizer (1). In Appendix B we will show that for any finite T, the process $\{(Z_t, \theta_t), t \leq T\}$ (with $\theta_t := \theta_{[\frac{t}{\mu}]}$) can be approximated by the solution of system of ODEs (3) and (5), ($R_{yy}(.)$, $R_{ys}(.)$ are defined in the Appendix B)

$$\dot{\theta}(t) = -R_{yy}(Z(t))\theta(t) + R_{ys}(Z(t)). \quad (5)$$

When $\eta < 0$ and $d_2 \in (-1, 1]$, i.e., when the channel is stable, it is easy to see that the channel ODE (3) has a unique global attractor. In the following Lemma we show that the equalizer ODE (5) also has a unique global attractor (the proof is provided in [4]).

Lemma 1: With $\eta < 0$ and $d_2 \in (-1, 1]$, if Z^* is the unique attractor of ODE (3) and θ^* is the Wiener filter corresponding to channel Z^*, then θ^* is the unique globally stable attractor for ODE (5).

IV. Performance Analysis

We will first show that a linear equalizer tracks the instantaneous Wiener filter of the channel, whenever the channel is not marginally stable. We study this performance using the solutions of the ODEs (3), (5) which approximate the trajectories of the channel and equalizer closely for small values of μ. We will also show that the difference between the instantaneous MSE of the linear equalizer and the instantaneous MMSE converges to zero.

Towards this end we first define the error process $E(t)$ as the error between the equalizer $\theta(t)$ and the Wiener filter corresponding to the channel value at t, i.e.,

$$\begin{aligned} E(t) &\stackrel{\triangle}{=} \theta(t) - \theta_*(t) \quad \text{where} \\ \theta_*(t) &\stackrel{\triangle}{=} [R_{yy}(Z(t))]^{-1} R_{ys}(Z(t)). \end{aligned}$$

We have proved the following theorem (see [4]).

Theorem 1: The process $E(t)$ decays exponentially with time:

$$|E(t)| \leq c_1 e^{-\sigma_n^2 t} + c_2 e^{-|a|t},$$

whenever $Z(t) = C_1 e^{at} + C_2$. It decays linearly/polynomially with time :

$$|E(t)| \leq c_1 e^{-\sigma_n^2 t} + \frac{c_2}{t^n},$$

whenever $Z(t) = C_1 t^n + C_2$, $n = 1$ or 2. □

Let $M(t)$ define the difference between the MSE and MMSE at time t, i.e.,

$$\begin{aligned} M(t) \stackrel{\triangle}{=} & \left[\theta(t) R_{yy}(Z(t))\theta(t) - R_{ys}(Z(t))\theta(t)\right] \\ & - \left[\theta_*(t) R_{yy}(Z(t))\theta_*(t) - R_{ys}(Z(t))\theta_*(t)\right]. \end{aligned}$$

We have proved that (details in [4]), the process $M(t)$ decays exponentially with time whenever $Z(t) = C_1 e^{at} + C_2$, with $a < 0$ (i.e., for stable channels) :

$$|M(t)| \leq c_1' e^{-\sigma_n^2 t} + c_2' e^{-|a|t}.$$

For unstable channels (exponentially / polynomially raising channels) we have shown that $M(t)$ does not explode with time, if required by using the Optimum Stepsize given below (details are in P.18, [4]).

When $d_2 = -1$ and $\eta < 0$ (a marginally stable case) the channel is approximated by a cosine waveform. Then we will see in Section V that the error may not decay to zero. For all other cases, the above theorem shows that the processes $E(t)$ reduces to zero with time.

Optimum Stepsize for the equalizer :

With $d_1 + d_2 \neq 1$ and $d_2 \neq -1$ (either stable or unstable channels), we see that the error decays exponentially at a rate, which is equal to the minimum of $\sigma_n^2, \; \frac{|d_1+d_2-1|}{\mu(1+d_2)}$. By introducing a constant r in the equalizer adaptation as below,

$$\theta_{k+1} \quad = \quad \theta_k - \mu r Y_k(\theta_k^T Y_k - s_k),$$

we can see that the error $E_r(t)$ (we have introduced the subscript r to show the dependence on factor r) decays as,

$$|E_r(t)| \leq c_1 exp^{-r\sigma_n^2 t} + c_2 exp^{-\frac{|d_1+d_2-1|}{\mu(1+d_2)}t}.$$

Now the new error decay rate is equal to minimum of $r\sigma_n^2, \; \frac{|d_1+d_2-1|}{\mu(1+d_2)}$. If $\sigma_n^2 << \frac{|d_1+d_2-1|}{\mu(1+d_2)}$, the error decay rate in Theorem 1, which equals minimum of $\{\sigma_n^2, \; \frac{|d_1+d_2-1|}{\mu(1+d_2)}\}$, reduces considerably. This can be compensated by choosing an appropriate scaling factor r. Hence one can choose optimum

$$r_* = max\left\{\frac{|d_1 + d_2 - 1|}{\sigma_n^2 \mu\,(1+d_2)}, 1\right\},$$

to achieve the fastest decay in the error. But please note that this decay only considers the first order analysis and hence if the value of r is very large such that the overall equalizer adaptation stepsize $r_*\mu$ is large then, one must choose a smaller r.

Similar optimum stepsize can be obtained even for unstable channels, given by $Z(t) = C_1 exp^{at} + C_2$ with $a > 0$. In Section V, in Figure 4, we have shown the improvement obtained, with $\mu_* = r_*\mu$ as stepsize.

Complex Signals and Channel :

The above analysis can be extended to a system with complex input signals and a complex channel. Towards this, we will have to convert each complex vector into a double dimensional real vector and then get the corresponding ODEs in the real domain. We have carried all these operations in Section V of [4] and have shown that the norms of the processes $E(t)$ and $M(t)$ behave in the same way as in the case of real signals.

V. Examples

In this section we illustrate the theory developed so far using some examples. We consider a three tap channel with a two tap equalizer. $W_k \sim \mathcal{N}(Mn, 1)$ for varying values of mean Mn. Input is BPSK.

In Figure 1, in the first subfigure, we plot the actual trajectory and the ODE solution of the AR(2) process. In the second subfigure of Figure 1, we plot the corresponding equalizer trajectory and the ODE solution for $\sigma_n^2 = 0.1$ and 0.5. We also, plot the instantaneous Wiener filter in the same figure. In Figure 2, we plot the same (the real and complex parts of the channel and equalizer coefficients),

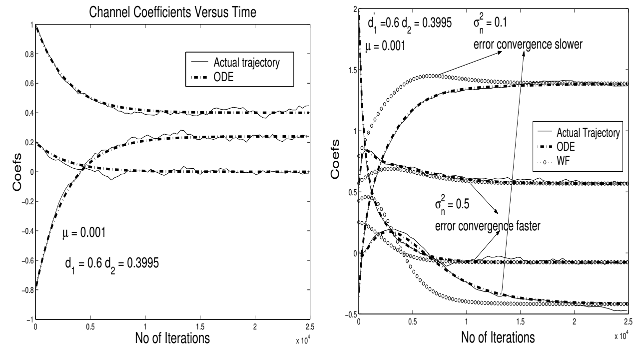

Fig. 1. Trajectories of AR(2) process (3 tap channel), Equalizer coefficients along with the trajectory of Wiener filter

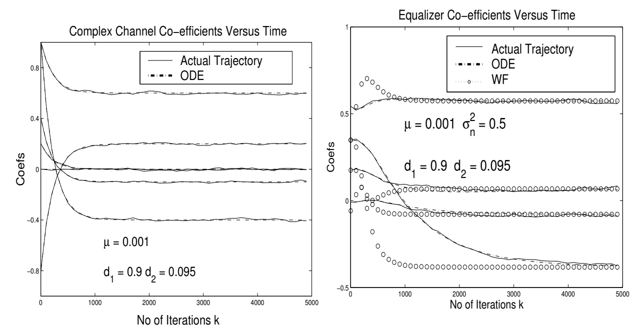

Fig. 2. Trajectories of Complex Channel and Equalizer coefficients

for a complex channel with BPSK input for $\sigma_n^2 = 0.5$. The channel parameters are provided in the two figures.

It is clear from both the figures, that the ODE solutions approximate the actual trajectories well. Figure 1 shows that the error $E(t)$ decays faster with $\sigma_n^2 = 0.5$. This is also evident from Theorem 1. But as one can see from Figure 1, this is mainly because the Wiener filter itself does not change much with time for the same channel if the noise variance is increased.

We have also varied d_1, d_2 values keeping Mn, η fixed and seen (via simulation results not provided here) that the decay rate increases with decrease in d_2, which is once again evident from Theorem 1. In simulations we have seen that as $1 - (d_1 + d_2)$ increases, the AR process converges faster to the attractor and this channel will be like a channel without drift. In this case, it is very close to a IID Gaussian random variable. This is also evident from the theory developed as in this case $|\eta|$ will be larger.

In Figure 4, we adapted the linear equalizer by optimum stepsize, $r_*\mu$, given in Section IV. We see that the linear equalizer catches up with the instantaneous Wiener filter faster even for small σ_n^2. Also, please note that in this figure we actually approximated the channel with an AR(4) process. We used the ODE approximation suggested in [4]. We see that the suggested ODE approximation is quite accurate even for an AR(4) process.

In Figure 3, we considered a marginally stable channel. Here the channel is approximated by an AR(2) process with $d_2 = -1$ and $d_1 < 2$. We see that these channels can be approximated by cosine waveforms as suggested by the theory developed. But one can see from Figure 3, that the error between the LMS linear equalizer and the instantaneous Wiener filter actually oscillates and does not converge to zero. Thus, we can conclude that for this kind of channels, the error may not reduce to zero with time. Please note that we used a bigger stepsize for the equalizer to achieve better tracking performance.

We plotted the instantaneous MSE versus time in Figure 5. We have plotted it for the channel of Figure 1 for two values of noise variances σ_n^2, 0.1 and 1.0. We can see from the figure that the MSE converges exponentially to the steady state value as is given by Theorem 1.

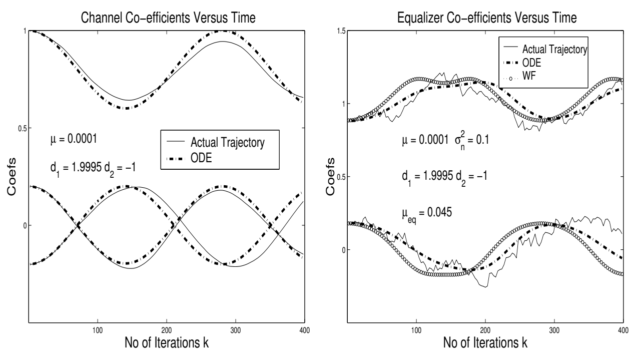

Fig. 3. Trajectories of AR(2) process (3 tap channel), Equalizer coefficients with $d_2 = -1$ and $d_1 < 2$.

VI. Conclusions and Remarks

In this paper we study an LMS-linear equalizer tracking an AR(2) process. We obtain the first order analysis of the error, which is defined as the deviation of the equalizer value from the instantaneous Wiener filter, using the ODE method. Towards this end we obtained the ODE of a general system whose components may depend upon 2 previous values. We then showed that the error between the equalizer and the Wiener filter falls exponentially/polynomially with time if the channel is not marginally stable. For the marginally stable case the error may not converge to zero. The MSE is also shown to approach the instantaneous MMSE exponentially with time when the channel is stable. It is also shown that the difference between the MSE and the MMSE does not explode with time even when the channel unstable.

This study can be extended to an AR(n) process even for $n > 2$. Also, we have only shown the theory when all the taps of the channel are given by the same AR(2) process. But in simulations we also considered different AR(2) processes for different channel taps. This extension in the theory is very easy but is not presented to make the explanations simple. In future, we will also be extending the above results to stationary non IID input signals. In this paper we have not considered the problem of designing an equalizer with delays greater than one. This is once again done to keep the discussions simple and the analysis can be easily extended.

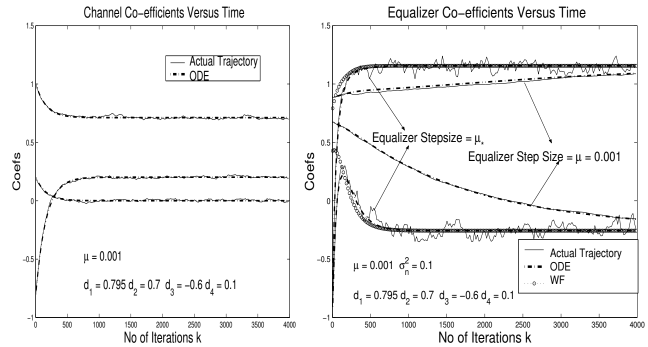

Fig. 4. Trajectories of AR(4) process, Equalizer coefficients. The improvement obtained with optimum Stepsize $= \mu_* = r_* \mu$.

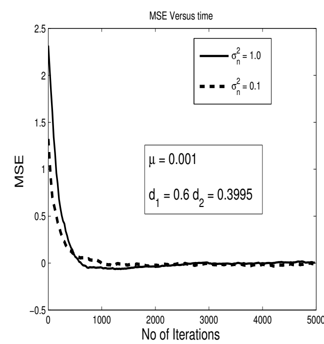

Fig. 5. MSE Versus Time

Appendix A : ODE Approximation for AR(2) process

In this Appendix we consider a general system given by,

$$Z_{k+1} = (1 - d_2) Z_k + d_2 Z_{k-1} + \mu H(W_k, Z_k). \tag{6}$$

We have shown in [4] that, this system can be approximated by ODE (with $h(Z(t)) := E\left[H\left(W_k, Z(t)\right)\right]$)

$$\begin{array}{rcll} \dot{Z}(t) & = & \frac{1}{1+d_2} h(Z(t)), & if \quad d_2 \in (-1, 1] \\ \frac{d^2 Z(t)}{dt^2} & = & h(Z(t)), & if \quad d_2 = -1, \end{array} \tag{7}$$

under the following assumptions.

A.1 W_k is an IID sequence.
A.2 $h(Z)$ is a C^1 function.
A.3 For any compact set Q there exist constants $C_1(Q), C(Q)$ such that,
$$E|H(Z, W)|^2 \leq C(Q) \ \ \forall Z \in Q.$$

A.3.i $E|H(Z, W)|^4 \leq C_1(Q) \ \ \forall Z \in Q.$

Let $Z(t, t_0, A)$ represent the solution of ODE (7) with initial condition $Z(t_0) = A$. With $d_2 = -1$, we require the additional initial condition $\dot{Z}(0) = 0$. Let Q_1 and Q_2 be any two compact sets such that $Q_1 \subset Q_2$ and we choose $T > 0$ such that there exists an $\delta_0 > 0$ and

$$d\left(Z(t, 0, A), Q_2^c\right) \geq \delta_0 \ \ \forall A \in Q_1 \text{ and } \forall t \leq T. \tag{8}$$

Theorem 2: Under the above Assumptions for all $\delta < \delta_0$, with $Z_0 = Z_{-1} = A$,

$$P\left\{\sup_{1\le k\le \lfloor T/\mu\rfloor} |Z_k - Z(k\mu,0,A)| \ge \delta\right\} \to 0 \quad if \quad d_2 \in (-1,1],$$

$$P\left\{\sup_{1\le k\le \lfloor T/\sqrt{\mu}\rfloor} |Z_k - Z(k\sqrt{\mu},0,A)| \ge \delta\right\} \to 0 \quad if \quad d_2 = -1,$$

as $\mu \to 0$ uniformly for all $A \in Q_1$.
Theorems 1 and 2 of [4] constitute the above theorem.

Now we return to the specific example of the above system, the AR(2) process (2). It can be written in the above general form by,

$$Z_{k+1} = (1-d_2)Z_k + d_2 + \mu(W_k + \eta Z_k)$$

where $\eta := \frac{d_1+d_2-1}{\mu}$. Clearly assumptions **A**.1-**A**.3 are satisfied for the AR(2) process. It is easy to see that condition (8) is satisfied by the ODE solution (4) for all T, when appropriate Q_1, Q_2 and $\delta_0 > 0$ are chosen. Thus the AR(2) process can be approximated by the ODE (3).

Appendix B : ODE Approximation for AR(2) and equalizer processes

Consider the following general system,

$$\begin{aligned} Z_{k+1} &= (1-d_2)Z_k + d_2 Z_{k-1} + \mu H(Z_k, W_k), \quad (9) \\ \theta_{k+1} &= \theta_k + \mu H_1(Z_k, \theta_k, X_{k+1}) \quad (10) \end{aligned}$$

where equation (9) satisfies all the conditions in Appendix A and the equation (10) satisfies the assumptions **B**.1 to **B**.4 given below. We will show that the above equations can be approximated by the solution of the ODE's (7) and (11).

$$\dot{\theta}(t) = h_1(Z(t), \theta(t)), \quad (11)$$

where $h_1(.,.)$ is defined in assumption **B**.3 We make the following assumptions, which are similar to that in [1]. We assume in the following that D is an open subset of $\mathcal{R}^d$.

B.1 There exists a family $\{\Pi_{Z,\theta}; \forall Z, \theta\}$ of transition probabilities $\Pi_{Z,\theta}(X, \mathbf{A})$ on $\mathcal{R}^k$ such that, for any Borel subset $\mathbf{A}$ of $\mathcal{R}^k$, we have

$$P[X_{n+1} \in \mathbf{A} | \mathcal{F}_n] = \Pi_{Z,\theta}(X_n, \mathbf{A})$$

where $\mathcal{F}_k \stackrel{\triangle}{=} \sigma(\theta_0, Z_0, Z_1, X_0, X_1, \cdots, X_k)$. This in turn implies that the tuple $(X_k, \theta_k, Z_k, Z_{k-1})$ will form a Markov chain.

B.2 For any compact set Q of D, there exist constants C_1, q_1 such that $\forall (Z,\theta) \in D$ and all n, we have

$$|H_1(Z,\theta,X)| \le C_1(1 + |X|^{q_1}).$$

B.3 There exists a function h_1 on D, and for each $Z, \theta \in D$ a function $\nu_{Z,\theta}(.)$ on $\mathcal{R}^k$ such that

a) h_1 is locally Lipschitz on D.

b) $(I - \Pi_{Z,\theta})\nu_{Z,\theta}(X) = H_1(Z,\theta,X) - h(Z,\theta)$.

c) For all compact subsets Q of D, there exist constants C_3, C_4, q_3, q_4 and $\lambda \in [0.5, 1]$, such that for all $Z, \theta, Z', \theta' \in Q$

$$\begin{aligned} |\nu_{Z,\theta}(X)| &\le C_3(1 + |X|^{q_3}), \\ |\Pi_{Z,\theta}\nu_{Z,\theta}(X) - \Pi_{Z',\theta'}\nu_{Z',\theta'}(X)| &\le C_4(1 + |X|^{q_4}) \\ &\quad \left|(Z,\theta) - (Z',\theta')\right|^{\lambda}. \end{aligned}$$

B.4 For any compact set Q in D and for any $q > 0$, there exists an $\mu_q(Q) < \infty$, such that for all $n, X \in \mathcal{R}^k$, $A = (Z,\theta) \in \mathcal{R}^d$

$$E_{X,A}\{I(Z_k, \theta_k \in Q, k \le n)(1 + |X_{n+1}|^q)\} \le \mu_q(Q)(1 + |X|^q)$$

where $E_{X,A}$ represents the expectation taken with $X_0 = X$ and $(Z_0, \theta_0) = A$.

Let $Z(t,t_0,Z), \theta(t,t_0,\theta)$ represent solutions of ODE's (7), (11) with initial conditions $Z(t_0) = Z$, $\theta(t_0) = \theta$. Once again for $d_2 = -1$, we need $\dot{Z}(0) = 0$. Let Q_1 and Q_2 be any two compact sets of D, such that $Q_1 \subset Q_2$ and we choose $T > 0$ such that there exists an $\delta_0 > 0$ and, for all $(Z,\theta) \in Q_1$, $t \le T$,

$$d\left((Z(t,0,Z), \theta(t,0,\theta)), Q_2^c\right) \ge \delta_0 \quad (12)$$

With $V \stackrel{\triangle}{=} (Z,\theta)$, representing the ordered pair Z, θ, we have proved Theorem 3 in [4], following the approach used in [1].

Theorem 3: With Assumptions **A**.1-**A**.3, **A**.3.i of Appendix A and **B**.1-**B**.4 of Appendix B, for all $\delta < \delta_0$ and for any initial condition X, with $Z_{-1} = Z_0 = Z$ and $\theta_0 = \theta$,

$$P_{X,Z,\theta}\left\{\sup_{1\le k\le \lfloor \frac{T}{\mu}\rfloor} |V_k - V(k\mu, 0, (Z,\theta))| \ge \delta\right\} \to 0$$

as $\mu \to 0$, uniformly for all $Z, \theta \in Q_1$ when $d_2 \in (-1,1]$.

When $d_2 = -1$, the same theorem follows, after replacing $V(k\mu, 0, (Z,\theta))$ in the statement of the above theorem with the ordered pair, $(Z(k\sqrt{\mu}, 0, Z), \theta(k\mu, 0, \theta))$(the details are given in [4]).

Now we verify that the conditions of Theorem (3) are satisfied by our system of Section III. Defining, $X_{k+1} \stackrel{\triangle}{=} [S_k^T, Y_k^T]^T$, one can rewrite the AR process and the equalizer adaptation as,

$$\begin{aligned} Z_{k+1} &= (1-d_2)Z_k + d_2 Z_{k-1} + \mu(W_k + \eta Z_k) \\ \theta_{k+1} &= \theta_k + \mu H_1(\theta_k, X_{k+1}) \\ H_1(\theta_k, X_{k+1}) &\stackrel{\triangle}{=} -Y_k(\theta_k^T Y_k - s_k). \end{aligned}$$

$\{X_k\}$ is a chain whose transition probabilities $\Pi_{Z_k}(X, \mathbf{A})$ are functions of Z_k alone. Thus condition **B**.1 is satisfied. **B**.2 is also satisfied as for any compact set Q and for any $\theta \in Q$,

$$|H_1(\theta, X)| \le 2\left[max\left\{1, \sup_{\theta\in Q}|\theta|\right\}\right](1 + |X|^2).$$

For any fixed Z, i.e., for a fixed channel, it is easy to see that $X_k(Z)$ has a stationary distribution given by

$$\begin{aligned} \Psi_Z([s_1, s_2, \cdots s_n] \times \mathbf{A_1}) &= Prob(S = [s_1, s_2, \cdots s_n]) \\ &\quad Prob(N \in \mathbf{A_1} - \pi_Z[s_1, s_2, \cdots s_n]^T) \end{aligned}$$

where π_Z is the $M \times M + L - 1$ length convolutional matrix formed from vector Z and S, N are the input and noise vectors of length $M + L - 1$, M respectively.

Define,

$$\begin{aligned}
R_{yy}(Z) &\triangleq E_Z\left[Y(Z)Y(Z)^T\right] = (\pi_Z\pi_Z^T + \sigma_n^2 I), \\
R_{ys}(Z) &\triangleq E_Z\left[Y(Z)s\right] = \pi_Z[1\ 0\cdots 0]^T, \\
h_1(Z,\theta) &\triangleq E_Z(H_1(\theta, X(Z)) = -R_{yy}(Z)\theta + R_{ys}(Z), \\
\nu_Z(X) &\triangleq \sum_{k\geq 0}\Pi_Z^k(H_1(\theta,X) - h_1(Z,\theta)).
\end{aligned}$$

Since h_1 is continuously differentiable, it is locally Lipschitz. Thus conditions **B**.3a, **B**.b are met. Also, since $\{X_k\}$ is a linear dynamical process depending upon the parameter Z, condition **B**.3c can be verified using Proposition 10 in Chapter 2, page 270 of [1] (The exact details are given in [4]).

The condition **B**.4 is trivially met as for any $n > M+L-1$, the expectation does not depend upon the initial condition X but is bounded based on the compact set Q and because of the Gaussian random variable N and discrete random variable S.

We have shown in [4] (Lemma 1 and Lemma 2) that the solution of the equalizer ODE (5) is bounded for any finite time T. Thus for any finite time T, one can choose appropriate Q_1, Q_2 and δ_0 such that condition (12) is satisfied. Thus all the conditions in Theorem 3 are met verifying the ODE approximation for the equalizer and the AR process of Section III.

References

[1] A. Benveniste, M. Mietivier, P. Priouret, "Adaptive Algorithms and Stochastic Approximation", Springer-Verlag, April 1990.

[2] M.Hajivandi, W.A. Gardner, "Measures of Tracking Performance of the LMS Algorithm", IEEE Trans. Acoust., Speech, Signal Processing, vol. 38, pp. 1953 1958, Nov. 1990.

[3] S. Haykin, "Adaptive Filter theory", Third Edition, Prentice-Hall International inc, 1996.

[4] V. Kavitha, V. Sharma, " Tracking performance of an LMS-Linear Equalizer for fading channels using ODE approach", Technical report noTR-PME-2005-11, DRDO-IISc programme on mathematical engineering, ECE Dept., IISc, Bangalore, October 2005. (downloadable from http://pal.ece.iisc.ernet.in/PAM/tech_rep05.html)

[5] C. Komminінakis, C.Fragouli, A. Sayed, R.Wesel, "Multiple-input Multiple-output Fading Channel Tracking and Equalization using Kalman Estimation", IEEE Trans. on Signal Process., Vol 50, No. 5, May 2002, pp 1064-1076.

[6] H. Kushner, G. George Yin, "Stochastic Approximation Algorithms and Applications", Springer, 1997.

[7] H. Kushner, J. Yang, "Analysis of adaptive step-size SA algorithms for parameter tracking", IEEE Trans. Automat. Contr., Vol. 40, pp. 1403 1410, Aug. 1995.

[8] L. Lindbom, M. Sternad and A. Ahlen, "Tracking of Time-Varying Mobile Radio Channels Part II: A Case Study", IEEE Trans. Communications, Vol. 50, No. 1, JANUARY 2002.

[9] L.Guo, L. Ljung, G-J Wang, "Necessary and Sufficient Conditions for Stability of LMS", IEEE Trans on Automatic control, Vol. 42, No. 6, June 1997, pp 761-770.

[10] O. Macchi, E. Eweda, "Convergence Analysis of Self-Adaptive Equalizers", IEEE Trans on Information Theory, Vol. 30, No. 2, March 1984, pp 161-176.

[11] D. P. Taylor, G.M. Vitetta, B.D. Hart, A. Mammela, "Wireless channel equalization", Eur. Trans. Telecom., Vol. 9,1998, 117-143.

[12] M. Tsatsanis, G. Giannakis, G. Zhou, "Estimation and Equalization of Fading Channels with Random Coefficients", Signal Process., Vol 53, no. 2-3, pp 211-229, Sept 1996.

[13] Wee-Peng Ang, B. Farhang-Boroujeny, "A New Class of Gradient Adaptive Step-Size LMS Algorithms", IEEE Trans. On Signal Processing, Vol. 49, No. 4, April 2001.

[14] B. Widrow, J. M. McCool, M. G. Larimore, C. R. Johnson, Jr., "Stationary and nonstationary learning characteristics of the LMS adaptive filter", Proc. IEEE, vol. 64, pp. 1151-1162, Aug. 1976.