# Analysis of an LMS Linear Equalizer for Fading Channels in Decision Directed mode

Veeraruna Kavitha and Vinod Sharma
Department of Electrical Communication Engineering,
Indian Institute of Science, Bangalore, 560012, India.
kavitha@pal.ece.iisc.ernet.in vinod@ece.iisc.ernet.in

***Abstract*— We consider a time varying wireless fading channel, equalized by an LMS linear equalizer in decision directed mode (DD-LMS-LE). We study how well this equalizer tracks the optimal Wiener equalizer. Initially we study a fixed channel. For a fixed channel, we obtain the existence of DD attractors near the Wiener filter at high SNRs using an ODE (Ordinary Differential Equation) approximating the DD-LMS-LE. We also show, via examples, that the DD attractors may not be close to the Wiener filters at low SNRs. Next we study a time varying fading channel modeled by an Auto-regressive (AR) process of order 2. The DD-LMS equalizer and the AR process are jointly approximated by the solution of a system of ODEs. We show via examples that the LMS equalizer ODE show tracks the ODE corresponding to the instantaneous Wiener filter when the SNR is high. This may not happen at low SNRs.**

*Key words:* Fading channels, LMS, Decision-Directed mode, 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.

An equalizer is most commonly designed using the Minimum Mean Square Error (MMSE) criterion ([6], [11], [17]). The optimal MSE (MMSE) equalizer, also called the Wiener filter (WF), is either calculated directly using the training sequence or indirectly using a training based channel estimate. The WF often involves a matrix inverse computation. Hence a computationally simpler iterative algorithm, the Least Mean Square (LMS), is commonly used as an alternative.

A Least Mean Square linear equalizer (LMS-LE), designed using training sequence, is a simple equalizer and is extensively used ([6], [11], [17]). For a fixed channel its convergence to the Wiener filter has been studied in [1], [13] (see also the references therein). For a time varying channel, theoretical tracking behavior (how well an LMS-LE tracks the instantaneous Wiener filter) has been studied in [8] (its tracking behavior is also studied via simulations, approximations and upper bounds on probability of error in [6], [10], [18]).

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 ([12], [19]). In fact, it is sufficient to model the fading channel by an AR(2) process ([10], [12], [19]). Thus, in [8] we model a time varying wireless channel by an AR(2) process. It is shown that for a stable/unstable channel (the poles are either inside/outside the unit circle) the LMS-LE tracks the instantaneous WF. It is also shown that for a marginally stable channel (the poles are on the unit circle), the distance between the LMS-LE and the instantaneous WF remains bounded.

A training based LMS-LE becomes inefficient in a wireless scenario. Due to time varying nature of the wireless channel, the training based LMS-LE, needs frequent transmission of the training sequence. Therefore, a significant ($\sim$ 18% in GSM) fraction of the channel capacity is consumed by the training sequence. The usual blind equalization techniques have also been found to be inadequate [5] due to their slow convergence and/or high computational complexity. In [7] it is shown using information theoretic arguments that a semi-blind method can be a better alternative for a time varying channel. In such scenarios, the decision directed LMS-LE (the training sequence is replaced by the decisions of the symbols after some time and hence such a DD-LMS-LE can also be viewed as a semi blind algorithm) may prove to be a good alternative ([13]).

However, one needs a theoretical understanding of the DD-LMS-LE prior to its use. In [13], it has been shown that the DD-LMS-LE for a fixed channel converges to the WF almost surely, if the initializer is sufficiently close to the WF. But the authors in [13] assume bounded channel output and perfect equalizability. These assumptions are not satisfied in most of the practical channels, e.g., an AWGN channel with ISI. In [13], the authors also deal with the AWGN noise and observe that the DD attractors are away from the WFs when the noise is non zero. However, they restrict themselves to a single tap equalizer. But ISI can be mitigated only with equalizers of length greater than one. The existence of undesirable local minima are established in [14], [15]. In [5] (Chapter 11 and the references therein) the convergence properties (noiseless) and initialization strategies (to 'open' eye) are discussed.

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

Hence, the DD-LMS-LE is not completely understood even for a fixed channel.

In this paper we first study a DD-LMS-LE on a fixed channel. We obtain an ODE approximation for its trajectory and show that the ODE's attractors are close to the corresponding Wiener filters as the noise variance tends to zero (whenever perfect equalizability is achieved at zero noise). We also show, once again using ODE approximation, that for large noise variances (i.e., at low SNRs) the DD attractors may not be close to the WFs. These results are obtained under more realistic conditions than in [13]. In particular, the equalizer can have more than one tap and the channel output need not be bounded. Furthermore, we assume perfect equalizability only at zero noise power.

Next we consider a DD-LMS-LE tracking a time varying wireless channel modeled by an AR(2) process. We use the ODE approximation of the AR(2) process (obtained by us in [8]) and obtain an ODE approximation for a DD-LMS-LE tracking an AR(2) process. Using this ODE approximation we illustrate via some examples that a DD-LMS-LE can indeed track an AR(2) process reasonably (the DD-LMS-LE trajectory is quite close to the instantaneous WFs) as long as the SNR is high. With increase in noise variance the DD algorithm loses out. We are not aware of any other theoretical study on the tracking behavior of a DD-LMS-LE.

The paper is organized as follows. In Section II we explain our model. Section III studies the decision directed (DD) algorithm on a fixed channel. Section IV obtains the ODE approximation for a time varying channel. Section V provides examples to demonstrate the ODE approximations and the proximity of the DD attractors to that of the WFs. Section VI concludes the paper. The appendices contain some details on the proofs.

## II. System Model, Notations and Assumptions

We consider a system consisting of a time varying (wireless) channel followed by an adaptive linear equalizer. The input of the channel $s_k$ comes from a finite alphabet and forms a zero mean IID (independent, identically distributed) process. The channel is a time varying finite impulse response (FIR) linear filter $\{Z_k\}$ of length $L$ followed by additive white Gaussian noise $\{n_k\}$. We assume $n_k \sim \mathcal{N}(0, \sigma^2)$. We also assume that $\{s_k\}$ and $\{n_k\}$ are independent of each other. The channel output at time $k$ is

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

where $Z_{k,i}$ is the $i^{th}$ component of $Z_k$. At the receiver the channel output $u_k$ passes through a linear equalizer $\theta_k$ and then through a hard decoder $Q$. The output of the hard decoder at time $k$ is $\hat{s}_k$.

In this paper we consider a DD-LMS-LE. For this system the LE $\theta_k$, of length $M$ at time $k$, is initially updated using a training sequence. After a while, the training sequence is replaced by the decisions made at the receiver about the current input symbol $s_k$. This is the decision directed (DD) mode.

The output $\hat{s}_k$ of the hard decoder $Q$ is $Q(\theta_k^t U_k)$, where $S_k, N_k, U_k$ are the appropriate length input, noise and channel output vectors respectively. We assume $E[S_k S_k^T] = I$. Note that, $U_k = \pi_k S_k + N_k$, where the convolutional matrix $\pi_k$ depends upon the channel co-efficients $Z_k \cdots Z_{k-M+1}$ and is given by,

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

In this paper we assume the input to be BPSK, i.e., $s_k \in \{+1, -1\}$. This assumption is made to simplify the discussions and can easily be extended to any finite alphabet source. For BPSK, $Q(x) = 1_{\{x>0\}} - 1_{\{x \le 0\}}$.

In DD mode the LE is updated using $\hat{s}_k(\theta)$: ,

$$\theta_{k+1} = \theta_k - \mu_k U_k (\theta_k^T U_k - \hat{s}_k(\theta_k)) \qquad (1)$$

where $\mu_k$ is a positive sequence of step-sizes.

Initially we study the DD system when the channel is fixed, i.e., $Z_k = Z$ for all $k$. Later on, we consider a time varying channel when the channel is modeled by an AR(2) process:

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

where $W_k$ is an IID sequence, independent of the processes $\{s_k\}, \{n_k\}$. An AR(2) process can approximate a wireless channel quite realistically ([10], [19]) and has been approximated by an ODE in [8]. Using this ODE approximation we obtain the required tracking performance analysis.

The fixed channel is studied in Section III while the time varying in Section IV.

## III. DD-LMS-LE for a fixed channel

In this section, we assume that the channel is fixed, i.e., $Z_k = Z$ for all $k$. We first obtain an ODE approximation for it when the step-sizes $\mu_k \to 0$. We obtain the existence of DD attractors (ODE) near the corresponding Wiener filters at high SNRs under the assumption of perfect equalizability for the channel with zero noise. We show that as noise variance $\sigma^2$ tends to zero, these DD attractors tend to the corresponding WFs.

### *A. ODE approximation*

DD-LMS-LE for a fixed channel has been approximated by an ODE in [1]. We start our analysis with this ODE. Towards this goal, as a first step the DD-LMS-LE algorithm (1) is rewritten to fit in the setup of [1], p. 276,

$$\begin{aligned} \xi_k &:= \begin{bmatrix} S_k^t & U_k^t & \hat{s}_k \end{bmatrix}^t, \\ H(\theta, \xi) &:= U^t \left(\theta^t U - \hat{s}\right), \\ \theta_k &= \theta_{k-1} - \mu_{k-1} H(\theta_{k-1}, \xi_k). \end{aligned}$$

Let $\theta(t, t_0, a)$ denote the solution of the following ODE with initial condition $\theta(t_0) = a$ ($\pi_Z$ is the convolutional matrix $\pi_k$ of the previous section for a fixed channel $Z$),

$$\begin{aligned}\dot{\theta}(t) &= -R_{uu}\theta(t) + R_{u\hat{s}}(\theta(t)),\\ R_{uu} &= \pi_Z \pi_Z^t + \sigma^2 I,\\ R_{u\hat{s}} &= E\left[UQ\left(U^t\theta\right)\right].\end{aligned}$$

It is easy to see that the Markov chain $\{\xi_k\}$ has a unique stationary distribution for every $\theta$ and that the DD-LMS satisfies all the required hypothesis of Theorem 13, p. 278, [1]. Hence one can approximate its trajectory on any finite time scale with the solution of the above ODE. We reproduce the precise result below.

For any initial condition $\theta_0$ and for any finite time T, with $t(r) := \sum_{k=0}^{r} \mu_k$, $m(n,T) := max_{r\geq n}\{t(r) - t(n) \leq T\}$

$$\sup_{\{n\leq r\leq m(n,T)\}} |\theta_r - \theta\left(t(r),\ t(n),\ \theta_0\right)| \xrightarrow{p} 0$$

as $n \to \infty$, whenever $\sum_k \mu_k = \infty$, $\sum_k \mu_k^{1+\delta} < \infty$ for some $\delta < 0.5$, $\mu_k \leq 1$ for all $k$ and $\liminf_k \frac{\mu_{k+r}}{k} > 0$ for all $r$. Also, from the above convergence one can easily see that the DD-LMS-LE trajectory converges to an attractor of the ODE in probability whenever the DD-LMS-LE is started in its region of attraction (see more details in [9]).

As in Lemma 1 of Appendix C one can show that, the above ODE has a unique global bounded solution for any finite time. We will also show the existence of attractors for this ODE, near WF, at least at high SNRs in the next subsection.

From the ODE approximation, if the decision directed mode of the system is started in the region of attraction of an attractor of the ODE, the DD-LMS-LE will converge to that attractor in probability. We will show below that under high SNR, an attractor of the above ODE will be close to the WF. Thus, the DD mode should be started when the LE is within the region of attraction of this attractor (e.g., when the 'eye' has opened as in [13]). To reach the region of attraction, one starts with a 'good' initial condition and then uses a training sequence. The region of attraction of a desired attractor depends upon the channel $Z$, the input distribution and $\sigma^2$. However, for a given set of parameters it may be computed via the various available methods ([4]).

### *B. Relation between DD attractors and WFs*

In the following we study the desired attractors in more detail.

Using implicit function theorem ([2]), we will show that the DD-LMS attractors are close to the WFs at high SNRs. Let (note that $R_{uu}$, $R_{u\hat{s}}$ depend on $\sigma^2$),

$$\begin{aligned}f(\theta,\sigma^2) &\stackrel{\triangle}{=} -R_{uu}\theta + R_{u\hat{s}}(\theta),\\ \theta^*(\sigma^2) &\stackrel{\triangle}{=} R_{uu}^{-1}R_{us},\ \text{where } R_{us} = E[Us].\end{aligned}$$

Note that $\theta^*(\sigma^2)$ represents the WF at noise variance $\sigma^2$, while a DD attractor is a zero of the function $f$. At $\sigma_n^2 = 0$, by invertibility of $\pi_Z\pi_Z^*$ and zero noise perfect equalizability (which we assume and this common assumption is for example discussed in Theorem 2 in [3]), $R_{u\hat{s}}(\theta^*(0))$ equals $R_{us}$. Hence $\theta^*(0)$, the WF at zero noise variance, also becomes a DD attractor. Thus, $(\theta^*(0), 0)$ is a zero of the function $f(.,.)$. One can easily verify the following :

- $f(\theta, 0) = -R_{uu}\theta + R_{us}$ whenever $\theta \in B(\theta^*(0), \epsilon)$ for some $\epsilon > 0$, where $B(x,r) = \{y : |x-y| \leq r\}$.
- Thus, $\frac{\partial f}{\partial\theta}(\theta^*(0), 0) = -R_{uu}$ and $R_{uu}$ is invertible.
- By Lemma 2, $f$ is continuously differentiable .

By implicit function theorem (Theorem 3.1.10, p. 115, [2]), there exists a $\sigma_0^2 > 0$ and a unique differentiable function $g$ of $\sigma^2$ such that, for all $0 \leq \sigma^2 \leq \sigma_0^2$,

$$f(g(\sigma^2), \sigma^2) = 0.$$

Since $\frac{\partial f}{\partial\theta}(\theta^*(0), 0) = -R_{uu}$ is negative definite and $\frac{\partial f}{\partial\theta}$ is continuous at $(\theta^*(0), 0)$, $\frac{\partial f}{\partial\theta}$ is negative definite on a small neighborhood around $(\theta^*(0), 0)$. Thus zeros, $g(\sigma^2)$ are DD attractors for all $\sigma^2$ small enough. We represent these DD attractors at noise variance $\sigma^2$, by $\theta_d^*(\sigma^2)$.

We will now relate the DD attractors, $\theta_d^*(\sigma^2) = g(\sigma^2)$, to the corresponding WFs, $\theta^*(\sigma^2)$ when $\sigma^2$ is close to zero. Define $h(\sigma^2) = R_{u\hat{s}}(\theta_d^*(\sigma^2))$. Using dominated convergence theorem and continuity of the map $g$, one can see that $h(\sigma^2) \to h(0) = R_{us}$ whenever $\sigma^2 \to 0$. Define

$$m(\theta, \sigma^2, \eta) = -R_{uu}\theta + R_{us} + \eta.$$

At any noise variance, $\sigma^2$, $m(\theta^*(\sigma^2), \sigma^2, 0) = 0$ as $\theta^*(\sigma^2)$ is the unique WF at noise variance $\sigma^2$. Also, the function $m$ is $C^\infty$ (infinitely differentiable) in all parameters (note that $R_{us}$ is a fixed vector independent of all the parameters whenever input is IID). Hence once again using implicit function theorem at any noise variance, $\sigma_0^2$ there exist $\alpha, \beta > 0$ and a continuous function $\gamma(.,.)$ such that,

$$m(\gamma(\sigma^2,\eta), \sigma^2, \eta) = 0 \text{ when } |\eta| \leq \beta, \text{ and } \left|\sigma^2 - \sigma_0^2\right| \leq \alpha.$$

Hence by continuity of the functions $\gamma$ and $h$, the WF (which is also given by $\gamma(\sigma^2, 0)$) will be close to the DD attractor, $\gamma\left(\sigma^2, \left[R_{us} - R_{u\hat{s}}(\theta_d^*(\sigma^2))\right]\right)$ at low noise variances.

## IV. DD-LMS-LE TRACKING AN AR(2) PROCESS

In this section we present the ODE approximation for the linear equalizer (1) in decision directed mode when the channel is modeled as an AR(2) process (2). Here we set the step-size $\mu_k = \mu$ for all $k$, to facilitate tracking of the time varying channel. We use Theorem 2 of Appendix B (parts of this Theorem are presented as Theorem 3 in [8]) to obtain the required ODE approximation.

We will show below that the trajectory $(\theta_k, Z_k)$ given by equations (1), (2) can be approximated by the solution of the following system of ODEs,

$$\begin{aligned}(1+d_2)\,\dot{Z}(t) &= \left[E(W) + \eta Z(t)\right], \quad \text{if } d_2 \in (-1,1],\\ \frac{d^2Z(t)}{dt^2} &= \left[E(W) + \eta Z(t)\right], \quad \text{if } d_2 = -1,\\ \frac{d^2Z(t)}{dt^2} + \eta_1\,\dot{Z}(t) &= \left[E(W) + \eta Z(t)\right],\\ &\qquad \text{if } d_2 \text{ is close to } -1, \qquad (3)\end{aligned}$$

$$\begin{aligned}
\dot{\theta}(t) &= -R_{uu}(Z(t))\theta(t) + R_{u\hat{s}}(\theta(t), Z(t)), \quad (4)\\
\eta &\triangleq \frac{d_1+d_2-1}{\mu}, \quad \eta_1 = \frac{1+d_2}{\sqrt{\mu}}\\
R_{uu}(Z) &\triangleq E_Z\left[U(Z)U(Z)^T\right] = (\pi_Z\pi_Z^T + \sigma^2 I),\\
R_{u\hat{s}}(\theta, Z) &\triangleq E_Z\left[U(Z)\hat{s}(\theta)\right].
\end{aligned}$$

In (3) , when $d_2$ is close but not equal to $-1$, two ODEs approximate the same AR(2) process. This is an important case and results when a second order AR process approximates a fading channel with a U-shaped band limited spectrum. It is obtained for small values of $f_dT$ where $f_d$ is the Doppler frequency shift and $T$ is the transmission time of one symbol. For example if $f_dT$ equals $0.04$, $0.01$ or $0.005$ the channel is approximated by an AR(2) process with $(d_1, d_2, \mu)$ equal to $(1.9707, -0.9916, 0.00035)$, $(1.9982, -0.9995, 1.38e^{-6})$ and $(1.9995, -0.9999, 8.66e^{-8})$ respectively (see, e.g., [10])). One could approximate such an AR(2) process with the first order ODE of (3). However this approximation will not be very accurate and will require $\mu$ to be very small. In this case, the second order ODE approximates the channel trajectory better. We will plot these approximations in Section V.

By Lemma 1, the above system of ODEs has unique bounded global solutions for any finite time. Let $Z(t,t_0,Z), \theta(t,t_0,\theta)$ represent the solutions of the ODEs (3), (4) with initial conditions $Z(t_0) = Z$, $\theta(t_0) = \theta$ and $\dot{Z}\ (t_0) = 0$ whenever the channel is approximated by a second order ODE.

Let $V_k \triangleq (Z_k, \theta_k)$ and $V(k) \triangleq (Z(\mu^\alpha k, 0, Z), \theta(\mu k, 0, \theta))$, where $\alpha = 1$ if $Z(.,.,.)$ is solution of a first order ODE and $1/2$ otherwise. We prove Theorem 1 using Theorem 2 of Appendix B.

**Theorem 1:** For any finite $T > 0$, for all $\delta > 0$ and for any initial condition $(G, \theta, Z)$, with $d_2 Z_{-1} + d_1 Z_0 = Z$, $\dot{Z}\ (t_0) = 0$ whenever the channel is approximated by a second order ODE and $\theta_0 = \theta$,

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

as $\mu \to 0$, uniformly for all $(Z, \theta) \in Q$, if $Q$ is contained in the bounded set containing the solution of the ODEs (3), (4) till time $T$.

**Proof :** Please see the Appendix A.

One can easily see that the solution of the channel (AR(2) process) ODE is,

$$Z(t) := \begin{cases}
C_1 e^{\frac{\eta}{1+d_2}t} - \frac{E(W)}{\eta}, & \eta \ne 0, d_2 \in (-1,1],\\
\frac{E(W)}{1+d_2}t + C_1, & \eta = 0, d_2 \in (-1,1],\\
C_1 cosh(\sqrt{\eta}\ \ t) - \frac{E(W)}{\eta}, & \eta > 0, d_2 = -1,\\
C_1 cos(\sqrt{|\eta|}\ \ t) - \frac{E(W)}{\eta}, & \eta < 0, d_2 = -1,\\
\frac{E(W)}{2}\ \ t^2 + C_1, & \eta = 0, d_2 = -1,\\
C_1 e^{-2at} + \frac{E(W)}{2a}\ \ t, & \eta = 0, d_2 \text{ close to } -1,\\
C_1 e^{-at} cos(\sqrt{|\eta| - a^2}\ \ t) - \frac{E(W)}{\eta}, & \eta < -a^2, d_2 \text{ close to } -1,\\
C_1 e^{-at} cosh(\sqrt{\eta + a^2}\ \ t) - \frac{E(W)}{\eta}, & otherwise,
\end{cases} \quad (5)$$

where the constant $C_1$ is chosen appropriately to match the initial condition of the approximated AR(2) process.

The approximating ODE (4) suggests that, its instantaneous attractors will be same as the DD-LMS-LE attractors obtained in the previous section when the channel is fixed at the instantaneous value of the channel ODE (3). We have shown in the previous section that these attractors are close to the WF at high SNRs. We will verify the same behavior for tracking, using some examples, in the next section.

One of the uses of the above ODE approximation is that, one can study the tracking behavior of the DD-LMS (e.g., proximity of its trajectory to the instantaneous WFs) using this ODE. This is done in the next section. Further, one can also obtain instantaneous theoretical performance measures like BER, MSE (approximately).

## V. Examples

In this section we illustrate the theory developed so far using some commonly used examples.

We first consider a fixed channel, $Z = [.41, .82, .41]$ in Figure 1. The channel of this example is very widely used (see p. 414, [6] and p. 165, [5]). We use a two tap linear equalizer. We plot the DD-LMS-LE, its ODE approximation and the Wiener filter for two values of noise variances $\sigma^2 = 0.01$, $1$ in this Figure. We can see that the ODE approximation is quite accurate for all the coefficients. We can also see that the DD-LMS coefficients as well as their ODE approximations converge to the DD attractor for both the noise variances. The ODE approximation thus confirms that with high probability the realizations of the DD-LMS trajectory (the DD-LMS trajectory in the figure being one such realization) converge to the attractor. One can see from this example that the DD-LMS attractors are close to the corresponding Wiener filters at high SNRs ($\sigma^2 = 0.01$) as is shown theoretically in Section III, but are away from the same at low SNRs ($\sigma^2 = 1$).

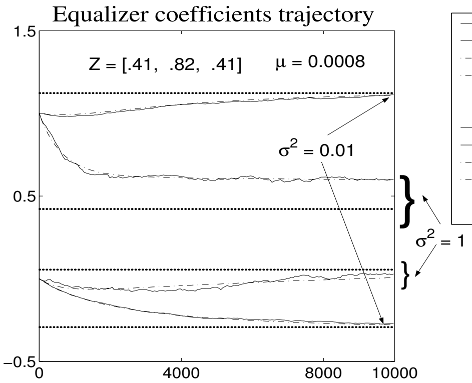


Fig. 1. The fixed channel equalizer coefficients for three tap channel $Z = [.41, .8, .41]$

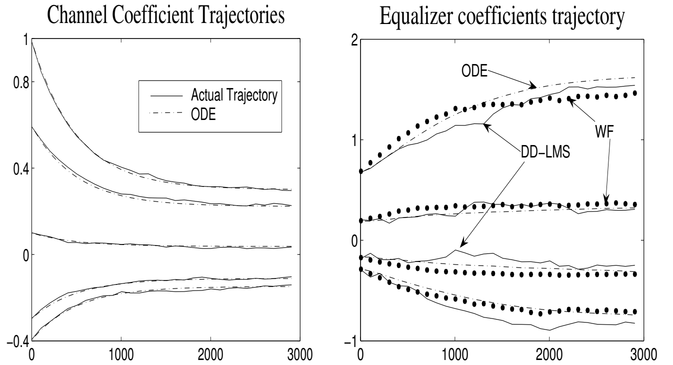


Fig. 2. Trajectories of AR(2) process, DD-LMS filter coefficients for a stable channel

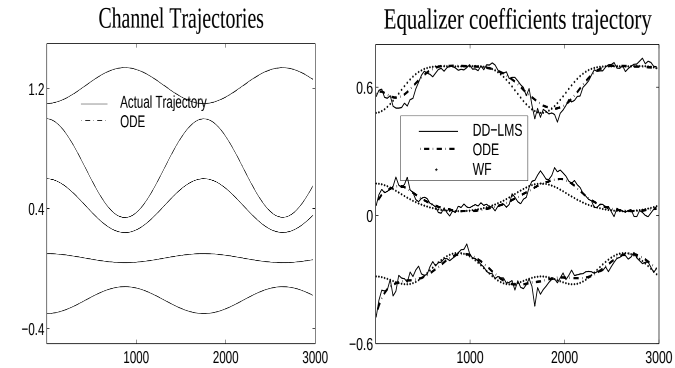


Fig. 3. Trajectories of AR(2) process, DD-LMS filter coefficients for a stable channel with $f_d T = 0.001$

We next consider two examples of a time varying channel equalized by a four/three tap equalizer. We consider two stable channels in Figures 2 and 3. In the second example corresponding to Figure 3, we consider a stable channel with $d_2$ close to $-1$. The AR parameters are $d_1 = .497$, $d_2 = 0.5$ and $\mu = 0.0007$ for the first stable channel, while the same parameters for the second stable channel are set at $1.999982$, $-.9999947$ and $1.399677e-010$ respectively. For the second channel, as is shown theoretically, a better ODE approximation is obtained by a second order ODE. Here, the channel trajectory is approximated by an exponentially reducing cosine waveform (as $d_2$ is very close to $-1$, the amplitude is reducing at a very small rate). This AR(2) process, approximates a fading channel with band limited and U-shaped spectrum and received with $f_d T = 0.001$. This can correspond, for example to a symbol time $T = 10\mu$s, at 2.4-GHz transmission with mobile speed of 45 Km/h .

Both the examples are run under high SNR conditions ($\sigma^2$ equals 0.05 for both the channels). In the first example, the DD-LMS and the ODE are started with the initial value of the WF while in the second they are started away from the initial WF. One can see from both the figures that the ODE once again approximates the DD-LMS quite accurately. Also, the DD-LMS and the ODE track the instantaneous WF quite well. Further, we can see from the second example that the DD-LMS and ODE catch up with the WF soon.

We further plot the instantaneous BER of the DD-LMS, the ODE and the WF in Figures 4, 5 respectively for both the stable channels. One can see that the performance of the DD-LMS and the ODE are quite close to that of the WFs throughout the time axis. The proximity of the ODE solution and the BER once again indicate that with high probability the realizations of DD-LMS track the instantaneous WFs.

Next we plot the DD-LMS, the ODE and the instantaneous WFs at two different noise variances in Figure 6 for a marginally stable channel. It is evident from the figure that the LMS-LE in DD mode, can track the channel variations at high SNR ($\sigma^2 = 0.05$), while it looses out at low SNRs ($\sigma^2 = 1$).

## VI. Conclusions

We have obtained theoretical performance analysis of an LMS linear equalizer in decision directed mode. We first studied a decision directed LMS-LE for a fixed channel. Using an ODE, which approximates the LMS-LE trajectory in decision directed mode, we showed the existence of DD attractors in the vicinity of the WFs at high SNRs. The same conclusion is also illustrated using some examples in Section V. Furthermore, we showed via examples that, a DD attractor may be away from the Wiener filter at low SNRs. We thus conclude that at high SNRs, one can update the LMS-LE in decision directed mode to obtain the WFs, by initializing it with a 'good' enough (the initializer must be in the region of attraction) training based estimate.

We next considered time varying channels. We modeled a time varying channel by an AR(2) process and obtained an ODE approximation for the tracking DD-LMS-LE. Using this ODE approximation, via some examples, we illustrated that LMS-LE in decision directed mode, can also be used to track the instantaneous WF at high SNRs. We also showed that, at low SNRs the decision directed mode does not provide a good equalizer.

We have extended these results to MIMO systems with complex input and channel parameters. (see [9]).

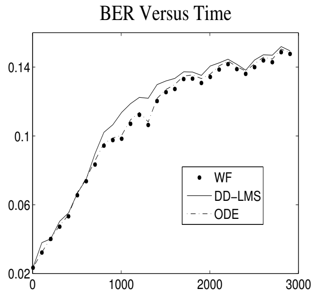


Fig. 4. Stable channel of Fig.2

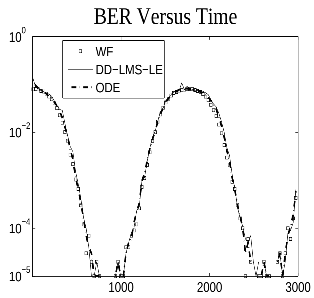


Fig. 5. Stable channel of Fig.3 with $f_dT = 0.001$

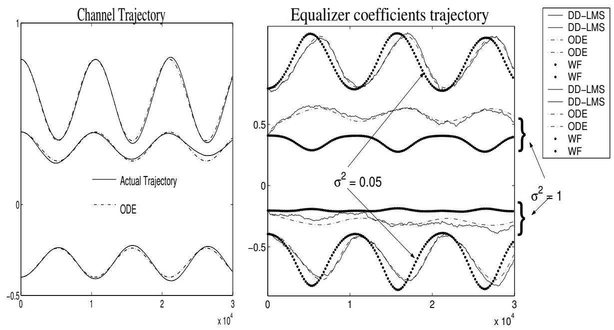


Fig. 6. DD-LMS versus WFs at varying $\sigma^2$ in a time varying channel

## APPENDIX A

In this appendix we prove Theorem 1 using Theorem 2 in Appendix B.

**Proof of Theorem 1 :** Defining $G_{k+1} \stackrel{\triangle}{=} [U_k^T, S_k^T]^T$, one can rewrite the AR process and the DD equalizer adaptation as,

$$\begin{aligned} Z_{k+1} &= (1-d_2)Z_k + d_2 Z_{k-1} + \mu(W_k + \eta Z_k) \\ \theta_{k+1} &= \theta_k + \mu H_1(\theta_k, G_{k+1}) \\ H_1(\theta_k, G_{k+1}) &\stackrel{\triangle}{=} -U_k(\theta_k^T U_k - \hat{s}_k) \\ &= -U_k(\theta_k^T U_k - Q(U_k^t \theta_k)). \end{aligned}$$

This is similar to the general system (6), (7) of Appendix B. Hence by Theorem 2, it suffices to show that our system satisfies the assumptions **A.**1-3, **B**.1-4 of Appendix B and that the above system of ODEs has a bounded solution for any finite time.

The AR(2) process $\{Z_k\}$ in (2) clearly satisfies the assumptions **A.**1 - **A.**3 as is shown in [8]. Assumption **B**.2 is satisfied as for any compact set $Q$ and for any $\theta \in Q$,

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

Fixing channel $Z_k = Z$ for all $k$, we obtain the transition kernel $\Pi_Z(.,.)$ for $\{G_k\}$ which is a function of $Z$ alone. Thus condition **B**.1 is satisfied. It is easy to see that $G_k(Z)$ 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 $\pi_Z$ is the $M \times M + L - 1$ length convolutional matrix formed from vector $Z$ (as in the fixed channel case) and $S, N$ are the input and noise vectors of length $M + L - 1$, $M$ respectively. Define,

$$\begin{aligned} h_1(\theta, Z) &\stackrel{\triangle}{=} E_Z(H_1(\theta, G(Z)) = -R_{uu}(Z)\theta + R_{u\hat{s}}(\theta, Z), \\ \nu_{\theta,Z}(G) &\stackrel{\triangle}{=} \sum_{k \ge 0} \Pi_Z^k(H_1(\theta, G) - h_1(Z, \theta)). \end{aligned}$$

By Lemma 2 of Appendix C, $h_1$ is locally Lipschitz. Thus conditions **B**.3 a, b are met.

We now prove condition **B**.3.c using Proposition 1 of [9] (which is proved as in [1]). $G_k$ is a linear dynamical process depending upon the channel realization $Z$ and it can be written as,

$$\begin{aligned} G_{k+1} &= A(Z)G_k + B(Z)\mathcal{W}_{k+1}, \text{ where,} \\ A(Z) &= \begin{bmatrix} \mathbf{J}_M & \mathbf{P} \\ \mathbf{0}_{L+M-1 \times L} & \mathbf{J}_{L+M-1} \end{bmatrix}, \; \mathcal{W}_{k+1} = [s_k, n_k], \\ B(Z) &= \begin{bmatrix} Z_0 & 1 \\ \multicolumn{2}{c}{\mathbf{0}_{M-1 \times 2}} \\ 1 & 0 \\ \multicolumn{2}{c}{\mathbf{0}_{L+M-2 \times 2}} \end{bmatrix}. \end{aligned}$$

In the above definitions, $\mathbf{0}_{n,m}$ is a $n \times m$ zero matrix. The matrix $\mathbf{J}_n$ is a $n \times n$ shift matrix, and $\mathbf{P}$ a $M \times L + M - 1$ matrix and these are given by,

$$\mathbf{J}_{n+1} = \begin{bmatrix} & 0_{1 \times n+1} \\ I_{n \times n} & 0_{n \times 1} \end{bmatrix}, \mathbf{P} = \begin{bmatrix} Z_1 & \cdots & Z_{L-1} & 0 & \cdots & 0 \\ & & \mathbf{0}_{M-1 \times L+M-1} & & & \end{bmatrix}.$$

It is easy to see that, $A^n(Z) = 0$ for all $n \ge max\{L, L + M - 1\}$ for all $Z$ as it involves the powers of shift matrices $\mathbf{J}_L$, $\mathbf{J}_{L+M-1}$, which satisfy $\mathbf{J}_n^n = 0$. By Lemma 3, the function $P_{\theta,Z}H(\theta, G)$ is $L_i(R^n)$ (see definition in [9]). Now all other conditions of Proposition 1, [9] are satisfied trivially (because $A(Z)$ and $B(Z)$ are linear in $Z$) and hence Proposition 1 holds and therefore, **B**.4.c holds for all $\lambda < 1$.

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

By Lemma 1 in Appendix C, the DD-LMS ODE has a unique bounded solution for any finite time. ■

## APPENDIX B : ODE APPROXIMATION OF A GENERAL SYSTEM

We consider the following general system,

$$\begin{aligned} Z_{k+1} &= (1-d_2)Z_k + d_2 Z_{k-1} + \mu H(Z_k, W_k), \quad (6) \\ \theta_{k+1} &= \theta_k + \mu H_1(Z_k, \theta_k, G_{k+1}), \quad (7) \end{aligned}$$

where equation (6) satisfies all the conditions in **A**.1–**A**.3 and the equation (7) satisfies the assumptions **B**.1–**B**.4, both given in the next para. We will show that the above equations can be approximated by the solution of the ODE's,

$$\begin{aligned}(1+d_2)\,\dot{Z}\,(t) &= h(Z(t)), \quad \text{if } d_2 \in (-1,1], \\ \frac{d^2 Z(t)}{dt^2} &= h(Z(t)), \quad \text{if } d_2 = -1, \\ \frac{d^2 Z(t)}{dt^2} + \eta_1\,\dot{Z}\,(t) &= h(Z(t)), \quad \text{if } d_2 \text{ is close to } -1,\end{aligned} \tag{8}$$

$$\dot{\theta}\,(t) = h_1(Z(t), \theta(t)), \tag{9}$$

where the function $h_1$ is defined in the assumptions given below and $h(Z) = E[H(Z,W)]$, $\eta_1 = \frac{1+d_2}{\sqrt{\mu}}$.

We make the following assumptions for the system (6) :

**A**.1 $\{W_k\}$ is an IID sequence.

**A**.2 $h(Z) = E\left[H\left(W_k, Z\right)\right]$ is a $C^1$ function.

**A**.3 For any compact set $Q$, there exists a constant $C(Q)$, such that $E|H(Z,W)|^2 \leq C(Q)$ for all $Z \in Q$, where the expectation is taken wrt $W$.

We make the following assumptions for (7), which are similar to that in [1]. Let $D \subset \mathcal{R}^d$ be an open subset.

**B**.1 There exists a family $\{P_{Z,\theta}\}$ of transition probabilities $P_{Z,\theta}(G, \mathbf{A})$ such that, for any Borel subset $\mathbf{A}$,

$$P[G_{n+1} \in \mathbf{A} | \mathcal{F}_n] = P_{Z_n,\theta_n}(G_n, \mathbf{A})$$

where $\mathcal{F}_k \stackrel{\triangle}{=} \sigma(\theta_0, Z_0, Z_1, W_1, W_2, \cdots, W_k, G_0, G_1, \cdots, G_k)$. This in turn implies that the tuple $(G_k, \theta_k, Z_k, Z_{k-1})$ forms a Markov chain.

**B**.2 For any compact subset $Q$ of $D$, there exist constants $C_1, q_1$ such that for all $(Z,\theta) \in D$ we have

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

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

a) $h_1$ is locally Lipschitz on $D$.

b) $(I - P_{Z,\theta})\nu_{Z,\theta}(G) = H_1(Z,\theta,G) - h_1(Z,\theta)$.

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 $Z, \theta, Z^{'}, \theta^{'} \in Q$

i) $|\nu_{Z,\theta}(G)| \leq C_3(1 + |G|^{q_3})$,

ii) $|P_{Z,\theta}\nu_{Z,\theta}(G) - P_{Z',\theta'}\nu_{Z',\theta'}(G)|$
$$\leq C_4\left(1 + |G|^{q_4}\right)\left|(Z,\theta) - (Z^{'},\theta^{'})\right|^{\lambda}.$$

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

$$\begin{aligned}E_{G,A}\left\{I(Z_k, \theta_k \in Q, k \leq n)\left(1 + |G_{n+1}|^q\right)\right\} \\ \leq \mu_q(Q)\left(1 + |G|^q\right),\end{aligned}$$

where $E_{G,A}$ represents the expectation taken with $G_0, Z_0, \theta_0 = G, Z, \theta$.

Let $Z(t, t_0, Z), \theta(t, t_0, \theta)$ represent the solutions of the ODEs (8), (9) with initial conditions $Z(t_0) = Z$, $\theta(t_0) = \theta$. For second order ODEs the additional initial condition is given by $\dot{Z}\,(t_0) = 0$. Let $Q_1$ and $Q_2$ be any two compact subsets of $D$, such that $Q_1 \subset Q_2$ and we can choose a $T > 0$ such that there exists an $\delta_0 > 0$ satisfying

$$d\left((Z(t,0,Z), \theta(t,0,\theta)), Q_2^c\right) \geq \delta_0, \tag{10}$$

for all $(Z,\theta) \in Q_1$ and all $t$, $0 \leq t \leq T$. We prove Theorem 2, following the approach used in [1]. Parts of this theorem are presented in [8].

**Theorem 2:** Assume, $E|H(Z,W)|^4 \leq C_1(Q)$ for all $Z$ in any given compact set $Q$ of $D$. Also assume **A**.1–**A**.3 and **B**.1–**B**.4. Furthermore, pick compact sets $Q_1$, $Q_2$, and positive constants $T$, $\delta_0$ satisfying (10). Then for all $\delta \leq \delta_0$ and for any initial condition $G$, with $Z_{-1} = Z_0 = Z$, $\dot{Z}\,(t_0) = 0$ (whenever $Z(.,.,.)$ is solution of a second order ODE), and $\theta_0 = \theta$,

$$P_{G,Z,\theta}\left\{\sup_{1 \leq k \leq \lfloor \frac{T}{\mu^\alpha} \rfloor} |(Z_k, \theta_k) - (Z(k\mu^\alpha, 0, Z), \theta(k\mu, 0, \theta))| \geq \delta\right\} \rightarrow 0 \text{ as } \mu \rightarrow 0$$

uniformly for all $Z, \theta \in Q_1$. If $Z(.,.,.)$ is solution of a first order ODE then $\alpha = 1$, otherwise $1/2$.

**Proof :** The proof is given in the Technical Report [9].

## APPENDIX C

**Lemma 1:** The ODE (4) has a unique solution which satisfies,

$$|\theta(t)| \leq c_0 + c_1 e^{-\sigma^2 t},$$

for appropriate positive constants $c_0$ and $c_1$.

**Proof :** For convenience, we reproduce the ODE (4),

$$\dot{\theta}\,(t) = -R_{uu}(Z(t))\theta(t) + R_{u\hat{s}}(\theta(t), Z(t)).$$

The matrix $R_{uu}(Z(t))$ is positive definite for all $t$, and it's minimum eigen value is greater than $\sigma^2$ for all $t$. Also, $|R_{u\hat{s}}(\theta(t), Z(t))| \leq C\,|Z(t)|$ for all $t$ for some constant $C > 0$. Using (5), $|R_{u\hat{s}}(\theta(t), Z(t))| \leq C(T)$ for all $t \leq T$ for any finite time $T$ for some positive constant $C(T)$ depending only on $T$. Thus, for any vector $\theta$, the inner product,

$$\begin{aligned}\left\langle \dot{\theta}\,(t), \theta \right\rangle &\leq -\sigma^2|\theta|^2 + C(T)|\theta| \\ &= \left[-\sigma^2|\theta| + C(T)\right]|\theta|.\end{aligned}$$

Therefore by Global existence theorem (pp 169 - 170 of [16]), the ODE (4), has a unique solution for any finite time and the solution is bounded by the solution of the scalar ODE (after choosing the initial conditions properly),

$$\dot{k}\,(t) = -\sigma^2 k(t) + C(T).$$

The solution of this ODE is $k(t) = c_1 e^{-\sigma^2 t} + C(T)$, for some appropriate constant $c_1$. ■

**Lemma 2:** The function $R_{u\hat{s}}(\theta, Z)$ is continuously differentiable in $(\theta, Z)$, $\sigma^2$ and hence is locally Lipschitz.

**Proof :** With $f_{\mathcal{N}}(\sigma^2, N)$ representing the $M$ dimensional Gaussian density with variance $\sigma^2$,

$$\begin{aligned} R_{u\hat{s}}(\theta, Z) &= E[(\pi_Z S + N)Q(\theta^t(\pi_Z S + N))] \\ &= \sum_S \int_{\{N:\theta^t(\pi_Z S+N)>0\}} (\pi_Z S + N) f_{\mathcal{N}}(\sigma^2, N) dN \\ &\quad - \int_{\{N:\theta^t(\pi_Z S+N)<0\}} (\pi_Z S + N) f_{\mathcal{N}}(\sigma^2, N) dN. \end{aligned}$$

We make the following change of variables,

$$\begin{aligned} Y &= A(\theta)(\pi_Z S + N) \text{ where matrix} \\ A(\theta) &\triangleq \begin{bmatrix} \theta_1 & \theta_2 & \cdots & \theta_M \\ 0 & 1 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & 1 \end{bmatrix}. \end{aligned}$$

With $|B|$ representing the determinant of the matrix B,

$$\begin{aligned} R_{u\hat{s}}(\theta, Z) &= \sum_S \int_{\{Y:Y_1>0\}} A(\theta)^{-1} Y \left| A(\theta)^{-1} \right| \\ &\qquad f_{\mathcal{N}}(\sigma^2, A(\theta)^{-1} Y - \pi_Z S) dY. \\ &\quad - \int_{\{Y:Y_1<0\}} A(\theta)^{-1} Y \left| A(\theta)^{-1} \right| \\ &\qquad f_{\mathcal{N}}(\sigma^2, A(\theta)^{-1} Y - \pi_Z S) dY, \end{aligned}$$

which is continuously differentiable by dominated convergence theorem and because the terms inside the integral are $C^\infty$. ■

Let $P_{\theta,Z}(.|.)$ represent the transition function of the Markov chain $G_k(\theta, Z)$ (when the channel and equalizer are fixed at $(\theta, Z)$).

**Lemma 3:** The function $P_{\theta,Z} H_\theta(G)$ is locally Lipschitz.

**Proof :** Note that,

$$\begin{aligned} P_{\theta,Z} H_\theta(G_0) &= E\left[H_1(\theta, G_1) | G_0 = (U_0, S_0)\right] \\ &= E\left[H_1\left(\theta, (A(Z)G_0 + B(Z)\mathcal{W}_1)\right)\right]. \end{aligned}$$

For all $(\theta, Z)$ in a compact set, one can get a positive constant $C_1$ depending only upon the compact set $Q$ such that,

$$\begin{aligned} &|P_{\theta,Z} H_\theta(G_0) - P_{\theta',Z'} H_{\theta'}(G_0')| \\ &\quad \leq E\left|(\theta^t U_1)U_1 - ({\theta'}^t U_1')U_1'\right| + 2E\left|U_1 - U_1'\right| \\ &\qquad + C_1 E\left|Q(\theta^t U_1) - Q({\theta'}^t U_1')\right|, \end{aligned}$$

where $U_1 \triangleq A(Z)G_0 + B(Z)\mathcal{W}_1$, $U_1' \triangleq A(Z')G_0' + B(Z')\mathcal{W}_1$. Suffices to show Lipschitz continuity for the last term. Now,

$$E\left|Q(\theta^t U_1) - Q({\theta'}^t U_1')\right| = 2P(Q(\theta^t U_1) \neq Q({\theta'}^t U_1')).$$

The Lemma follows because, using the steps as in Lemma 2, we can show that the above term is continuously differentiable. ■

## References


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

[2] M. S. Berger, "Nonlinearity and Functional Analysis", Academic Press, New York, 1977.

[3] Z Ding and L Qiu, "Properties of matrix polynomials and MIMO channel identifiability from second order statistics", IEEE Conference on decision and control, 1998.

[4] R. Genesio, M. Tartaglia and A. Vicino, "On the estimation of asymptotic stability regions: State of the art and new proposals", IEEE Transactions on Automatic Control, 1985, 747-755.

[5] G. B. Giannakis, Y. Hua, P. Stoica and L. Tong, "Signal Processing Advances in Wireless and Mobile Communications", Trends in Channel estimation and equalization. Vol.1, Prentice Hall, Upper Saddle River, NJ 2000.

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

[7] V. Kavitha and V. Sharma, " Comparison of training, blind and semiblind equalizers in MIMO fading systems using capacity as measure", ICASSP 05, USA.

[8] V. Kavitha and V. Sharma, "Tracking performance of an LMS-Linear Equalizer for fading channels", 44th Annual Allerton Conference on Communication, Control and Computing, USA, September 2006.

[9] V. Kavitha and V. Sharma, "DD/Training-based Linear Equalizers for MIMO Fading channels: Tracking Performance", Technical report no: TR-PME-2007-01, DRDO-IISc program on mathematical engineering, ECE Dept., IISc, Bangalore, January 2007. (downloadable from http://pal.ece.iisc.ernet.in/PAM/tech_rep07.html)

[10] C. Komminiakis, C.Fragouli, A. Sayed and 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.

[11] E. A. Lee and D. G. Messerschmitt, "Digital communications", 2nd Edn., Kluwer Academic Publishers, 1994.

[12] 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, Jan 2002.

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

[14] J. E. Mazo, "Analysis of decision-directed equalizer convergence", Bell Sysr. Tech. Journal, vol. 59, no. 10, pp. 1857-1876, Dec. 1980.

[15] G. Picchi and G. Prati, "Blind equalization and carrier recovery using a "stop-and-go" decision-directed algorithm", IEEE Trans. Commun., com-35, pp. 877-887, Sept. 1987.

[16] L.C. Piccinini, G. Stampacchia and G. Vidossich, "Ordinary Differential Equations in $R^n$", Springer-Verlag, NewYork, Volume 39, 1978.

[17] J. G. Proakis, "Digital Communications", New York: McGraw-Hill, 2000.

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

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