Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

Least-Square Estimation for Nonlinear Systems with Applications to Phase and Envelope Estimation in Wireless Fading Channels

Yiannis Socratous, Charalambos D. Charalambous, and Costas N. Georghiades

Abstract— **This paper addresses the problem of nonlinear Least-Square estimation. A new approach is presented which employees a change of probability measure technique to derive recursive equations for conditional means of nonlinear problems. The theory developed is applied to multipath fading wireless channels to derive phase and envelope estimates. Furthermore, some numerical results are presented in order to evaluate the performance of these estimators.**

I. INTRODUCTION

In classical Least-Square estimation one is interested in finding the best estimate, $\Phi^*(Y)$, of X from the measurements Y, by minimizing the expected value of the least-square error $e(X,\Phi) \stackrel{\triangle}{=} ||X - \Phi(Y)||^2_{\Re^n}$ over all functions $\Phi : \Re^d \mapsto \Re^n, Y \mapsto \Phi(Y)$, which is a function of Y ($||x||^2_{\Re^n}$ denotes Euclidean norm of $x \in \Re^n$). By the orthogonal projection theorem, the solution is given by the conditional expectation:

$$\Phi^*(Y) = E[X|Y] = \int_{\Re^n} x dP_{X|Y}(x|y) \qquad \text{(I.1)}$$

where $P_{X|Y}$ is the conditional distribution of X given Y [1], [2].

The solution to this estimation problem is stated in terms of the a posteriori density function $P_{X|Y}$. This density contains all information available about the state X. The objective is thus to estimate, recursively in time, the a posteriori density and its associated features. However, it is only in a few special cases that this density can be parameterized using a Finite number of statistics. The most important example is the case with linear dynamics and observations in additive Gaussian noise. In this case all involved densities are Gaussian, and hence they can be parameterized using the corresponding mean and covariance. The equations of the finite statistics are given by the Kalman Filter [1], [3].

In the case of nonlinear systems, there are serious difficulties in obtaining the solution of the a posteriori density in closed form. Often, approximations have to be made to find sub-optimal recursive nonlinear estimators. The standard approximation is to use the Taylor series expansion and apply linear filtering theory, giving rise to the so-called extended Kalman filter (EKF) [4], [5], [6]. Other more sophisticated estimation techniques than the EKF are available, as well, e.g., reiteration, higher order filters, and statistical linearization [4]. Although more advanced techniques generally improve the estimation accuracy, they involve complex implementation while they increase the computational burden.

The current paper, builds on the theoretical background of [1] to investigate nonlinear estimation problems for a class of systems, which include nonlinear recursive equations in discrete-time and autoregressive channel models found in [7], [8]. The methodology is based on nonlinear estimation theory using the concept of a "sufficient statistic". The "sufficient statistic" being the unnormalized a posteriori distribution, which is shown to satisfy a recursive equation. The recursive equations for the a posteriori distribution are derived using the change of probability measure technique [9]. The conditional distribution is a sufficient statistic in the sense that it conveys all the information of the observed sample path.

The theory developed is applied to multipath fading wireless channels. Typically, in a wireless fading channel, the a priori distribution of the channel parameters (phase, attenuation, etc.) are not Gaussian, and therefore linear filtering techniques can not be used to find the optimal Least-Squares estimates. This problem is a nonlinear estimation problem which can be solve using the recursive equation of the conditional distribution.

The rest of the paper is organized as follows. In Section II the general state and observation model is presented. In Section III the mathematical theory is presented and the recursive equation of the unnormalized a posteriori distribution is derived. In Section IV the theory of Section III is applied to multipath fading wireless channels. Various estimators are derived by solving the recursive equation satisfied by the unnormalized version of the a posteriori density. Finally, some numerical results are presented in order to evaluate the performance of the derived estimators.

II. STATE AND OBSERVATION MODELS

Let $\left(\Omega, \mathcal{F}, P\right)$ be a complete probability space on which the state or unobserved process $\{x_k\}, k \in N_+ \stackrel{\triangle}{=} \{0,1,2,3,\dots\}$ and the observation process $\{y_k\}, k \in N_+$, are defined by the following recursions:

$$\begin{aligned} x_{k+1} &= f(k+1, x_k) + B_{k+1} w_{k+1}, \quad x_0 \in \Re^n \\ y_k &= h(k, x_k) + D_k v_k, \quad y_0 \in \Re^d \end{aligned} \qquad \text{(II.2)}$$

Here $x_0 : \Omega \rightarrow \Re^n$ is the initial state and $w : \Omega \times N_+ \rightarrow \Re^n$, $v : \Omega \times N_+ \rightarrow \Re^d$, are random noises.

Also $\{w_k\}, \{v_k\}, k \in N_+$ are independent noise sequences of random variables with densities $\Psi_{w_k}(w) =$

Y. Socratous and C.D. Charalambous are with the Department of Electrical and Computer Engineering, University of Cyprus, Nicosia, CYPRUS e-mail: yianniss@ucy.ac.cy, chadcha@ucy.ac.cy.

C.N. Georghiades is with the Department of Electrical and Computer Engineering, Texas A&M University, Texas, USA e-mail: c-georghiades@tamu.edu.

$\frac{1}{(2\pi)^{\frac{n}{2}}}\ e^{-\frac{w^{tr}w}{2}}$, $\Xi_{v_k}(v) = \frac{1}{(2\pi)^{\frac{d}{2}}}\ e^{-\frac{v^{tr}v}{2}}$, respectively, x_0 has density $\pi_{x_0}(x) = \frac{d\Pi_{x_0}(x)}{dx}$, which is also independent of $\{w_k\}$, $\{v_k\}$.

We shall assume throughout that $f : N_+ \times \Re^n \rightarrow \Re^n$ and $h : N_+ \times \Re^d \rightarrow \Re^d$, B_k, D_k are Borel measurable functions, and $(B_k B_k^{tr})^{-1}$, $(D_k D_k^{tr})^{-1}$ exist (the importance of these conditions is discussed in [9]).

For background material on probability concept we use [9]. Let $\{\mathcal{G}_m^o\}$ be the σ-field generated by the complete data $\{x_0, x_1, \ldots, x_m, y_0, y_1, \ldots, y_m\}$ and let $\{\mathcal{G}_m\}, m \in N_+$ denote its complete filtration [9]. Let $\{\mathcal{Y}_m^o\}$ define the σ-field generated by the incomplete data $\{y_0, y_1, \ldots, y_m\}$ and let $\{\mathcal{Y}_m\}, m \in N_+$ denote its complete filtration. Let y^m denote the sequence $\{y_0, \ldots, y_m\}$ and similarly for other sequences. We denote by $\tilde{x}_m$ the estimate of the state x_m given $\{\mathcal{Y}_m\}, m \in N_+$; we assume recursive estimates which update $\tilde{x}_m$ from knowledge of $\tilde{x}_{m-1}$ and past and present data $\{y_0, \ldots, y_m\}$.

III. Recursive Equation for the Unnormilized Conditional Density

The next theorem presents intermediate steps using the change of probability measure discussed extensively in [9].

Theorem 3.1: Let Φ be a bounded continuous function on $\Re^n$ taking values in $\Re$.

Define the likelihood function of the complete data $\{x_0, \ldots, x_m, y_0, \ldots, y_m\}$ by

$\Lambda_m \triangleq \prod_{k=0}^m \left[\frac{\Xi_{v_k}(D_k^{-1}(y_k - h(k, x_k)))}{|D_k|\Xi_{v_k}(y_k)} \frac{\Psi_{w_k}(B_k^{-1}(x_k - f(k, x_{k-1})))}{|B_k|\Psi_{w_k}(x_k)} \right]$
$= \frac{dP(x^m, y^m)}{d\overline{P}(x^m, y^m)}$

where under probability distribution $\overline{P}$, $\{x_k\}$ is i.i.d. $N(0, I_n)$ and $\{y_k\}$ is i.i.d. $N(0, I_d)$ with density functions $\Psi_{w_k}(x_k) = \frac{1}{(2\pi)^{\frac{n}{2}}} \exp(\frac{-x_k^{tr} x_k}{2})$ and $\Xi_{v_k}(y_k) = \frac{1}{(2\pi)^{\frac{d}{2}}} \exp(\frac{-y_k^{tr} y_k}{2})$, $k \in N_+$, respectively.

Then we have the following results.

1) Conditional expectations are related via

$$E\Big[\Phi(x_m)\Big|\mathcal{Y}_m\Big] = \frac{\overline{E}\Big[\Phi(x_m)\Lambda_m|\mathcal{Y}_m\Big]}{\overline{E}\Big[\Lambda_m|\mathcal{Y}_m\Big]} \qquad \text{(III.3)}$$

where $\overline{E}$ denotes expectation under the probability distribution $\overline{P}$.

Moreover, the numerator of (III.3) can be written in terms of unnormalized probability distribution $\alpha_m(.)$ via

$$\alpha_m(\Phi) \triangleq \overline{E}\Big[\Phi(x_m)\Lambda_m|\mathcal{Y}_m\Big] = \int_{\Re^n} \Phi(z) d\alpha_m(z) \quad \text{(III.4)}$$

2) If $\alpha_m(\Phi)$ has a density, that is $\frac{d}{dx}\alpha_m(x) = \overline{\alpha}_m(x)$, then

$$\alpha_m(\Phi) = \int_{\Re^n} \Phi(x) d\alpha_m(x) = \int_{\Re^n} \Phi(x)\overline{\alpha}_m(x) dx \quad \text{(III.5)}$$

and the density $\{\overline{\alpha}_m(x)\}_{m \geq 0}$ satisfies the recursion

$$\overline{\alpha}_m(x) = \frac{\Xi_{v_m}(D_m^{-1}(y_m - h(m, x)))}{|D_m|\Xi_{v_m}(y_m)} \times \int_{\Re^n} \frac{\Psi_{w_m}(B_m^{-1}(x - f(m, z)))}{|B_m|}\ \overline{\alpha}_{m-1}(z) dz \qquad \text{(III.6)}$$

with initial condition

$$\overline{\alpha}_0(x) = \frac{\Xi_{v_0}(D_0^{-1}(y_0 - h(0, x)))}{|D_0|\Xi_{v_0}(y_0)}\ \pi_{x_0}(x) \qquad \text{(III.7)}$$

Proof. The derivation can be found in [9].

IV. Phase and Envelope Estimation for Frequency Selective Multipath Fading Channels

Consider a frequency-selective fading channel given by

$$\begin{aligned} y(t_k) &= \sum_{i=1}^N h_i(t_k, \theta_i, r_i) + D(t_k)v(t_k) \\ &= h(t_k, \boldsymbol{\theta}, \mathbf{r}) + D(t_k)v(t_k) \end{aligned} \qquad \text{(IV.8)}$$

where ω_c is the carrier frequency, $\{\tau_i(t_k)\}$ denotes the propagation delay, $\{r_i\}$, $\{\theta_i\}$ are random variables denoting the attenuation and phase, respectively, of the signal received associated with ith path, and $v(t_k) \sim N(0, 1)$. Define the phase vector $\boldsymbol{\theta} \triangleq (\theta_1, \ldots, \theta_N)'$, and attenuation vector $\mathbf{r} \triangleq (r_1, \ldots, r_N)'$, and $h_i(t_k, \theta_i, r_i) \triangleq r_i \cos(\omega_c(t_k - \tau_i(t_k)) + \theta_i)S(t_k - \tau_i(t_k))$.

Assume the delays $\{\tau_i(\omega, .)\}_{i=1}^N$ are fixed, while the phases $\theta_i : \Omega \rightarrow [0, 2\pi]$ are independent and identically distributed random variables with a priori density $\pi_{\theta_0}(\theta_i) = \frac{1}{2\pi}$, $\theta_i \in [0, 2\pi]$, while the attenuations $r_i : \Omega \rightarrow [0, \infty)$ are independent and identically distributed random variables with a priori density $\pi_{r_0}(r_i)$, for $1 \leq i \leq N$. In addition we assume $\{r_i\}_{i=1}^N$ and $\{\theta_i\}_{i=1}^N$ are independent, and also independent of the noise process $\{v(t_k); k \in N_+\}$.

The relation of model (IV.8) with model (II.2) is the following. Clearly, sampling time k of previous section is now represented by t_k, and the state vector of (IV.8) which needs to be estimated is $x = \begin{pmatrix} \boldsymbol{\theta} \\ \mathbf{r} \end{pmatrix} \in \Re^{2N}$. However, since $\begin{pmatrix} \boldsymbol{\theta} \\ \mathbf{r} \end{pmatrix}$ are random variables then $x_{k+1} = x_k$, $\forall k \in N_+$.

The density (III.6) is specialized to model (IV.8). This is done by replacing the integrand $\Psi_{w_t}(B_t^{-1}(x - f(t, z)))$, $x \in \Re^{2N}$, $z \in \Re^{2N}$ in (III.6) with the delta measure $\delta(\mathbf{r} - z^1) \times \delta(\boldsymbol{\theta} - z^2) = \prod_{i=1}^N \delta(r_i - z_i^1) \times \delta(\theta_i - z_i^2)$, where $x = \begin{pmatrix} \boldsymbol{\theta} \\ \mathbf{r} \end{pmatrix}$, $z = \begin{pmatrix} z^1 \\ z^2 \end{pmatrix} \in \Re^{2N}$, $f(t, z) = \begin{pmatrix} z^1 \\ z^2 \end{pmatrix}$ to get

$$\overline{\alpha}_m(\boldsymbol{\theta}, \mathbf{r}) = \pi_{\boldsymbol{\theta}_0}(\boldsymbol{\theta})\pi_{\mathbf{r_0}}(\mathbf{r}) \times \prod_{k=0}^m \left[\frac{\Xi_{v_k}(D^{-1}(t_k)(y(t_k) - \sum_{i=1}^N h_i(t_k, \theta_i, r_i)))}{|D(t_k)|\Xi_{v_k}(y(t_k))} \right] \qquad \text{(IV.9)}$$

where $\pi_{\boldsymbol{\theta}_0}(\boldsymbol{\theta}) = \prod_{i=1}^N \pi_{\theta_0}(\theta_i)$ is the a priori joint density of $(\theta_1, \ldots, \theta_N)$, and $\pi_{\mathbf{r_0}}(\mathbf{r}) = \prod_{i=1}^N \pi_{r_0}(r_i)$ is the a priori joint density of $(r_1, \ldots, r_N)$.

For the rest of this paper, we shall neglect the $2\omega_c$ terms ("double-frequency" terms) since the receiver will remove them in the process of reconstructing the transmitted signal. Furthermore, we also assume that the received paths are resolvable, meaning that inter-path delays are larger than the reciprocal of bandwidth of the transmitted signals. Therefore, under the resolvability assumption, the expressions containing cross terms are zero.

Using the above assumptions, and by further manipulation of (IV.9) we get a simplified version of the unnormalized conditional density

$$\overline{\alpha}_m(\boldsymbol{\theta}, \mathbf{r}) = \prod_{i=1}^N \Big[\pi_{\theta_0}(\theta_i)\pi_{r_0}(r_i) \exp(-r_i^2 K_i^m) \times \exp\Big(r_i V_i(y^m) \cos(\theta_i - \gamma_i(y^m)) \Big) \Big] \times \exp\Big(\frac{1}{2}\sum_{k=0}^m y^2(t_k)[1 - D^{-2}(t_k)] - \sum_{k=0}^m \log|D(t_k)| \Big) \quad \text{(IV.10)}$$

where $K_i^m = \frac{1}{4}\sum_{k=0}^m D^{-2}(t_k) S^2(t_k - \tau_i(t_k))$.

The following quantities are also needed when presenting the estimators.

Definition 4.1: For each $1 \le i, j \le N, i \ne j$, define

$V_c^i(y^m) \triangleq \sum_{k=0}^m D^{-2}(t_k) \cos(\omega_c(t_k - \tau_i(t_k))) S(t_k - \tau_i(t_k)) y(t_k)$

$V_s^i(y^m) \triangleq \sum_{k=0}^m D^{-2}(t_k) \sin(\omega_c(t_k - \tau_i(t_k))) S(t_k - \tau_i(t_k)) y(t_k)$

$V_i(y^m) = \sqrt{V_c^i(y^m)^2 + V_s^i(y^m)^2}$,

$\gamma_i(y^m) = -\tan^{-1}\left(\frac{V_s^i(y^m)}{V_c^i(y^m)} \right)$

Note that path resolvability implies that each path component (θ_i, r_i) can be estimated independently of the rest.

Theorem 4.1: Assuming the phases θ_i are independent and identically distributed random variables with a priori density $\pi_{\theta_0}(\theta_i) = \frac{1}{2\pi}$, $\theta_i \in [0, 2\pi]$ and the attenuations r_i are independent and identically distributed random variables with a priori density $\pi_{r_0}(r_i) = \frac{r_i}{\sigma^2} \exp -\frac{r_i^2}{2\sigma^2}$, $r_i \in [0, \infty)$ (Rayleigh distributed), then we obtain the following estimators.

(a) The incomplete data likelihood ratio defined by $\hat{\Lambda}(t_m) = \overline{E}[\Lambda(t_m)|\mathcal{Y}_m]$ is given by

$$\hat{\Lambda}(t_m) = \alpha_m(1) = \int_{[0,\infty)^n} \int_{[0,2\pi]^n} \overline{\alpha}_m(\boldsymbol{\theta}, \mathbf{r}) d\boldsymbol{\theta} d\mathbf{r} \quad \text{(IV.11)}$$

$$= \prod_{i=1}^N \left[\frac{1}{1 + 2\sigma^2 K_i^m} \exp\Big(\frac{V_i^2(y^m)\sigma^2}{2 + 4\sigma^2 K_i^m} \Big) \right] \times \exp\Big(\frac{1}{2}\sum_{k=0}^m y^2(t_k)[1 - D^{-2}(t_k)] - \sum_{k=0}^m \log|D(t_k)| \Big) \quad \text{(IV.12)}$$

where $I_0(.)$ is the modified bessel function of the first kind and zeroth order defined by
$I_0(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \exp(x\cos\alpha) d\alpha$
(b) The normalized conditional density of $(\boldsymbol{\theta}, \mathbf{r})$ given $\mathcal{Y}_m$, i.e., $p_N(t_m, \boldsymbol{\theta}, \mathbf{r}|\mathcal{Y}_m)$, is given by

$$p_N(t_m, \boldsymbol{\theta}, \mathbf{r}|\mathcal{Y}_m) = \frac{\overline{\alpha}_m(\boldsymbol{\theta}, \mathbf{r})}{\int\int \overline{\alpha}_m(\boldsymbol{\theta}, \mathbf{r}) d\boldsymbol{\theta} d\mathbf{r}} = \prod_{i=1}^N p_N(t_m, \theta_i, r_i|\mathcal{Y}_m) \quad \text{(IV.13)}$$

$$= \prod_{i=1}^N \left[\frac{r_i(1 + 2\sigma^2 K_i^m)}{2\pi\sigma^2} \exp\Big(-\frac{r_i^2(1 + 2\sigma^2 K_i^m)}{2\sigma^2} \Big) \times \exp\Big(r_i V_i(y^m) \cos(\theta_i - \gamma_i(y^m)) \Big) \times \exp\Big(-\frac{V_i^2(y^m)\sigma^2}{2 + 4\sigma^2 K_i^m} \Big) \right] \quad \text{(IV.14)}$$

(c) The minimum least-square estimator of θ_i given $\mathcal{Y}_m$ is given by

$$\tilde{\theta}_i^*(t_m) = E\Big[\theta_i|\mathcal{Y}_m\Big] = \int_0^\infty \int_0^{2\pi} \theta_i p_N(t_m, \theta_i, r_i) d\theta_i dr_i \quad \text{(IV.15)}$$

$$= \frac{(1 + 2\sigma^2 K_i^m)}{2\pi\sigma^2} \exp\Big(-\frac{V_i^2(y^m)\sigma^2}{2 + 4\sigma^2 K_i^m} \Big) \times \int_0^\infty \int_0^{2\pi} \theta_i \; r_i \exp\Big(-\frac{r_i^2(1 + 2\sigma^2 K_i^m)}{2\sigma^2} \Big) \times \exp\Big(r_i V_i(y^m) \cos(\theta_i - \gamma_i(y^m)) \Big) d\theta_i dr_i \quad \text{(IV.16)}$$

(d) The minimum least-square estimator of r_i given $\mathcal{Y}_m$ is given by

$$\tilde{r}_i^*(t_m) = E\Big[r_i|\mathcal{Y}_m\Big] = \int_0^{2\pi} \int_0^\infty r_i p_N(t_m, \theta_i, r_i) dr_i d\theta_i \quad \text{(IV.17)}$$

$$= (\sqrt{\pi/2}) \Big(\sqrt{\frac{\sigma^2}{1 + 2\sigma^2 K_i^m}} \Big) \exp\Big(-\frac{V_i^2(y^m)\sigma^2}{2 + 4\sigma^2 K_i^m} \Big) \times {}_1F_1\Big(\frac{3}{2}, 1; \frac{V_i^2(y^m)\sigma^2}{2 + 4\sigma^2 K_i^m} \Big) \quad \text{(IV.18)}$$

where ${}_1F_1\Big(\alpha, \beta; x\Big)$ is the confluent hypergeometric function [10].

Proof. The derivations can be found in [11].

A. Numerical Results and Discussion

We consider a multipath fading channel, given by (IV.8), with 3 arriving paths. Next, we assume that the phases θ_i are iid random variables with a priori density $\pi_{\theta_0}(\theta_i) = \frac{1}{2\pi}$, $\theta_i \in [0, 2\pi]$ and the attenuations r_i are also iid random variables with a priori density $\pi_{r_0}(r_i) = \frac{r_i}{\sigma^2_{r_i}} \exp -\frac{r_i^2}{2\sigma^2_{r_i}}$, $r_i \in [0, \infty)$ (Rayleigh distributed). We take the attenuation parameter σ_{r_i} for the three arriving paths as $[\frac{1}{\sqrt{2}}, 0.565685, 0.452548]$ respectively. It is also assumed that the time delay $\tau_i(t_k)$ for each path, is known precisely. Specifically, the time delay $\tau_i(t_k)$ is constant over time and is taken as $[0, 35, 70]$ $msec$ for each arriving path respectively.

Assume frequency $f_c = 1KHz$, signaling period $T_s = 30msec$, and a sampling time $\Delta_t = 10^{-4}sec$, transmitted signal $S(t_k) = 1$, $D(t_k) = 1$, and a Gaussian noice $v(t_k) \sim N(0, \sigma_n^2)$. Finally, the Signal to Noise Ratio (SNR) is defined as $SNR = \frac{P_s}{\sigma_n^2}$, were P_s is the power of the transmitted signal.

We are going to evaluate the performance of our estimators using the Mean Square Error (MSE) of each estimator for each path over time, for SNR=10 dB, and taking N=100 realizations. The general expression of the MSE is given by

$$MSE_{Z_i} = \frac{1}{N} \sum_{j=1}^{N} |Z_i - \tilde{Z}^*_{i,j}|^2 \qquad \text{(IV.19)}$$

where N is number of realizations, Z_i is the real value of the parameter we would like to estimate for path i and $\tilde{Z}^*_{i,j}$ is the estimated value of that parameter for path i.

First, we begin with the phase estimation. Figure (VI.1) displays the performance of the phase estimator (IV.16) over time for each arriving path for SNR=10 dB. We notice that the MSE decreases with time for all three paths, as time increases. This shows that our estimated phase converges to the real phase, as time increases, for all three paths.

Next, we continue with the estimation of the channel attenuation. Figure (VI.2) displays the performance of the attenuation estimator (IV.18) over time for each arriving path and for SNR=10 dB. Our observations, are similar with the ones experienced for the phase MSE. We can see the the MSE decreases as time increases.

V. Conclusion

This paper considers the problem of nonlinear estimation and employs a change of probability measure technique to derive recursive equations for conditional means. The theory developed is applied to frequency selective fading channels and the results derived include new estimators for phase and envelope estimation problems. Numerical examples are presented to evaluate the performance of these estimators. Future work should concentrate on the application of the theory to MIMO Channels.

VI. Acknowledgments

The research leading to these results has received funding from the Cyprus Research Promotion Foundation under the project $APTEMI\Sigma$.

References

[1] H. V. Poor, "An Introduction to Signal Detection and Estimation," New York: Springer-Verlag, 1994.

[2] H. L. Van.Trees, "Detection, Estimation and Modulation Theory-Part1," New York: Wiley, 1968.

[3] T. Kailath and H. Vincent Poor, "Detection of Stochastic Processes," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2230-2259, October 1998.

[4] A. Jazwinski, "Stochastic Processes and Filtering Theory," San Diego, CA: Academic, 1970.

[5] S. Haykin, "Kalman Filtering and Neural Networks," New York: Wiley, 2001.

[6] B. Anderson and J. Moore, "Optimal Filtering," Englewood Cliffs, NJ: Prentince-Hall, 1979.

[7] T. Feng, T. R. Field and S. Haykin, "Stochastic Differential Equation Theory Applied to Wireless Channels," *IEEE Trans. Commun.*, vol. 55, no. 8, pp. 1478-1483, August 2007.

[8] K. E. Baddour and N. C. Beaulieu, "Autoregressive Modeling for Fading Channel Simulation," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1650-1662, July 2005.

[9] R.J. Elliott, L. Aggoun and J.B. Moore, "Hidden Markov Models: Estimation and Control, " Springer-Verlag, Berlin-Heidelberg-New York 29,1994.

[10] J. G. Proakis, "Digital Communications," New York: McGraw-Hill, 4th ed., 2001.

[11] Y. Socratous, C. D. Charalambous and C. N. Georghiades, "Nonlinear Estimation with Applications to Phase and Envelope Estimation in Frequency Selective Wireless Fading Channels," (under preparation).

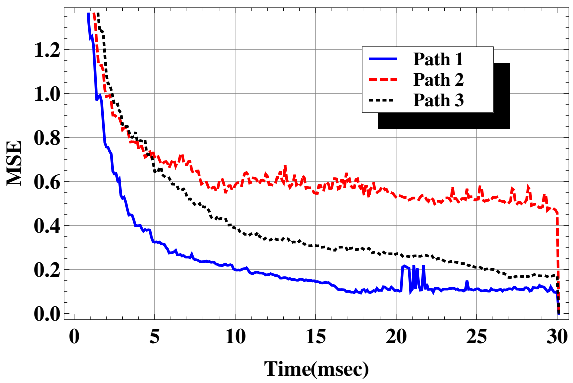

Fig. VI.1. MSE of the phase estimator, for SNR=10 dB

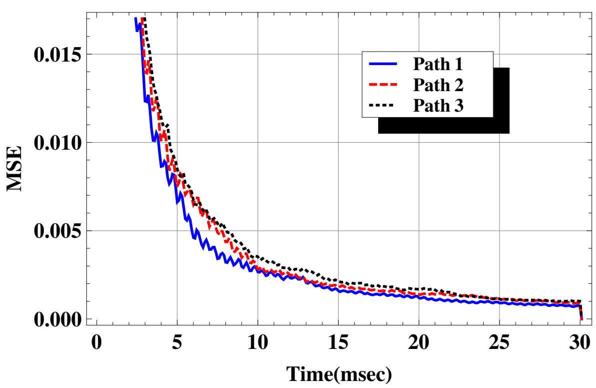

Fig. VI.2. MSE of the attenuation estimator, for SNR=10 dB