DESIGN OF THE CONSTRAINED ADAPTIVE LMS L-FILTERS

C. Kotropoulos I. Pitas

Department of Electrical Engineering
University of Thessaloniki
Thessaloniki 540 06, GREECE

ABSTRACT Two novel adaptive nonlinear filter structures are proposed which are based on linear combinations of order statistics. These adaptive schemes are modifications of the standard LMS algorithm and have the ability to incorporate constraints imposed on coefficients in order to permit location-invariant and unbiased estimation of a constant signal in the presence of additive white noise. The convergence in the mean of the filter coefficients is proven. The proposed filters can adapt well to a variety of noise probability distributions ranging from the short-tailed ones to long-tailed ones.

1 Introduction

Adaptive filters constitute an important part of statistical signal processing. They have been applied in a wide variety of problems [1]. An effort has been attempted to combine the adaptive filtering and the nonlinear filtering [2]. Extentions of the LMS and RLS algorithms have been proposed in [3]. Adaptive hybrid filter structures have been proposed in [4].

The main purpose of this paper is to extend the standard LMS algorithm by applying it to the adaptation of the coefficients of the L-filters in order to incorporate constraints imposed on the coefficients. L-filters are defined as linear combinations of the ordered data in the filter window, i.e. the output of the L-filter at time instant k is given by:

$$y(k) = \sum_{j=1}^{M} a_j \, x^k_{(j)} \tag{1}$$

where $x^k_{(j)}$ is the i-th largest observed data and M is assumed to be odd. It has been proven [5,6] that the optimal L-filter for estimating a constant signal in the presence of additive white noise should be either location-invariant or unbiased. Let $\mathbf{e}_M$ denotes the M×1 unitary vector, i.e., $\mathbf{e}_M = [1,\ldots,1]^T$ and $\mathbf{a} = [a_1,\ldots,a_M]^T$ denotes the vector of L-filter coefficients. The necessary and sufficient condition for a location-invariant L-filter is:

$$\mathbf{e}_M^T \mathbf{a} = 1 \tag{2}$$

The sufficient conditions for an unbiased L-filter are:

$$\mathbf{e}_M^T \mathbf{a} = 1$$
$$a_j = a_{M-j+1} \quad j=1,\ldots,(M-1)/2 \tag{3}$$

and the noise distribution should be symmetric about zero.

Two novel schemes are derived by rewriting the normal equations in a form that takes into account the constraints and by using instantaneous values for the correlations of the ordered noise samples in order to derive an estimate for the gradient vector.

2 Constrained LMS adaptive L-filters

A constant signal s corrupted by zero-mean additive white noise is considered. Thus, the input samples have the form $x_i = s + n_i$, where n_i are i.i.d. random variables having zero mean. The noise distribution is assumed symmetric about zero.

First, the adaptation formula for the location-invariant adaptive LMS L-filter is derived. Let $\mathbf{n}_r$ denote the vector of the ordered noise samples, i.e.:

$$\mathbf{n}_r = [n_{(1)}, n_{(2)}, \ldots, n_{(M)}]^T \tag{4}$$

The mean squared error J is given by:

$$J = E[\,(y(k)-s)^2\,] = \mathbf{a}^T \mathbf{R}\, \mathbf{a} \tag{5}$$

where $\mathbf{R}$ is the correlation matrix of the ordered noise samples. Its (i,j) element is given by $r_{ij} = E[n_{(i)} n_{(j)}]$, i,j =1,...,M. Let $\mathbf{e}$ denotes the (M-1)/2×1 unitary vector. The coefficient vector $\mathbf{a}$ is rewritten as:

$$\mathbf{a} = [\mathbf{a}_1^T \mid 1-\mathbf{e}^T\mathbf{a}_1-\mathbf{e}^T\mathbf{a}_2 \mid \mathbf{a}_2^T]^T \tag{6}$$

where $\mathbf{a}_1=[a_1,\ldots,a_{(M-1)/2}]^T$ and $\mathbf{a}_2=[a_{(M+3)/2},\ldots,a_M]^T$. Similarly, $\mathbf{n}_r$ can be partitioned in the form (6), i.e.:

$$\mathbf{n}_r = [\mathbf{n}_{r1}^T \mid n_{((M+1)/2)} \mid \mathbf{n}_{r2}^T]^T \tag{7}$$

where $n_{((M+1)/2)}$ is the median noise sample, $\mathbf{n}_{r1}=[n_{(1)}\ldots n_{((M-1)/2)}]^T$ and $\mathbf{n}_{r2}=[n_{((M+3)/2)},\ldots,n_{(M)}]^T$. Then, the correlation matrix $\mathbf{R}$ is partitioned as follows:

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{r}_1 & \mathbf{R}_2 \\ \mathbf{r}_1^T & r & \mathbf{r}_2^T \\ \mathbf{R}_3 & \mathbf{r}_2 & \mathbf{R}_4 \end{bmatrix} \tag{8}$$

where $\mathbf{R}_1= E[\mathbf{n}_{r1}\mathbf{n}_{r1}^T]$, $\mathbf{R}_2= E[\mathbf{n}_{r1}\mathbf{n}_{r2}^T]$, $\mathbf{R}_3= \mathbf{R}_2^T$, $\mathbf{R}_4= E[\mathbf{n}_{r2}\mathbf{n}_{r2}^T]$, $\mathbf{r}_1= E[n_{((M+1)/2)}\mathbf{n}_{r1}]$, $\mathbf{r}_2= E[n_{((M+1)/2)}\mathbf{n}_{r2}]$ and $r= E[n^2_{((M+1)/2)}]$. The MSE (5) is rewritten as:

$$J = r - 2\,\tilde{\mathbf{a}}^T \mathbf{p}' + \tilde{\mathbf{a}}^T \mathbf{R}' \tilde{\mathbf{a}} \tag{9}$$

where

$$\tilde{\mathbf{a}} = [\mathbf{a}_1^T \mid \mathbf{a}_2^T]^T$$

$$\mathbf{p}' = [r\mathbf{e}^T-\mathbf{r}_1^T \mid r\mathbf{e}^T-\mathbf{r}_2^T]^T \tag{10}$$

$$\mathbf{R}'= \begin{bmatrix} \mathbf{R}_1 + r\mathbf{e}\mathbf{e}^T - 2\mathbf{r}_1\mathbf{e}^T & \mathbf{R}_2 + r\mathbf{e}\mathbf{e}^T - 2\mathbf{r}_1\mathbf{e}^T \\ \mathbf{R}_3 + r\mathbf{e}\mathbf{e}^T - 2\mathbf{r}_2\mathbf{e}^T & \mathbf{R}_4 + r\mathbf{e}\mathbf{e}^T - 2\mathbf{r}_2\mathbf{e}^T \end{bmatrix}$$

The main advantage of this version of the MSE is that it leads to an adaptation scheme which does not use any heuristic technique to impose location-invariance e.g. the normalization of the coefficients that would be derived by a direct application of the LMS algorithm to the minimization of the MSE given in (5). The steepest descent algorithm for the minimization of J in (9) is given by:

$$\tilde{\mathbf{a}}(k+1) = \tilde{\mathbf{a}}(k) + \mu\,[\mathbf{p}' - \mathbf{R}_s'\,\tilde{\mathbf{a}}(k)] \tag{11}$$

where μ is the adaptation step and $\mathbf{R}_s'$ is the symmetric part of the matrix $\mathbf{R}'$. The bracketed term is the gradient $\nabla J(k)$ of MSE with respect to $\tilde{\mathbf{a}}(k)$. In the following, we shall drop the primes from $\mathbf{p}'$ and $\mathbf{R}_s'$. The simplest way to develop an estimate of the gradient $\nabla J(k)$ is to use instantaneous estimates for $\mathbf{p}$ and $\mathbf{R}_s$:

$$\hat{\mathbf{p}}(k) = n^k_{((M+1)/2)}\left(n^k_{((M+1)/2)}\mathbf{e}_{M-1} - \tilde{\mathbf{n}}_r(k) \right)$$
$$\hat{\mathbf{R}}_s(k) = \tilde{\mathbf{n}}_r(k)\,\tilde{\mathbf{n}}_r^T(k) - n^k_{((M+1)/2)}\mathbf{e}_{M-1}\tilde{\mathbf{n}}_r^T(k) + n^k_{((M+1)/2)}\left(n^k_{((M+1)/2)}\mathbf{e}_{M-1} - \tilde{\mathbf{n}}_r(k) \right)\mathbf{e}_{M-1}^T \tag{12}$$

(13)

The LMS adaptation formula is written as follows:

$$\hat{\mathbf{a}}(k+1) = \hat{\mathbf{a}}(k) + \mu\,[\hat{\mathbf{p}}(k) - \hat{\mathbf{R}}_s(k)\,\hat{\mathbf{a}}(k)] =$$
$$= \hat{\mathbf{a}}(k) + \mu\,\varepsilon(k)\,\left(\tilde{\mathbf{x}}_r(k) - x^k_{((M+1)/2)}\mathbf{e}_{M-1} \right) \tag{14}$$

where $\varepsilon(k) = s - y(k)$, $\tilde{\mathbf{x}}_r(k) = s\,\mathbf{e}_{M-1} + \tilde{\mathbf{n}}_r(k)$ and $x_{((M+1)/2)} = s + n_{((M+1)/2)}$. The coefficient for the median sample is given by:

$$\hat{a}_{(M+1)/2}(k) = 1 - \mathbf{e}_{M-1}^T\,\hat{\mathbf{a}}(k) \tag{15}$$

The structure of the proposed adaptive filter is shown in Figure 1.

We proceed next to the derivation of the unbiased LMS adaptive L-filter. Let $\mathbf{L}$ be the following $[(M-1)/2 \times (M-1)/2]$ matrix:

$$\mathbf{L} = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & & & \vdots \\ 1 & \cdots & 0 & 0 \end{bmatrix} \tag{16}$$

By using (3), the coefficient vector takes the following form:

$$\mathbf{a} = [\mathbf{a}_1^T \mid a_{(M+1)/2} \mid \mathbf{a}_1^T\,\mathbf{L}]^T \tag{17}$$

where $a_{(M+1)/2}= 1 - 2\,\mathbf{e}^T\mathbf{a}_1$. The correlation matrix of the ordered noise samples exhibits a double symmetry which is expressed by the equations:

$$\mathbf{r}_2 = \mathbf{L}\,\mathbf{r}_1,\ \mathbf{R}_3 = \mathbf{R}_2^T,\ \mathbf{R}_1 = \mathbf{L}\,\mathbf{R}_4\,\mathbf{L},\ \mathbf{R}_2\mathbf{L} = \mathbf{L}\,\mathbf{R}_3 \tag{18}$$

By employing (18), the MSE is given by the following expression:

$$J = r - 4\,\mathbf{a}_1^T\,\mathcal{D} + 2\,\mathbf{a}_1^T\,\mathcal{R}\,\mathbf{a}_1 \tag{19}$$

where

$$\mathcal{D} = r\mathbf{e} - \mathbf{r}_1 \tag{20}$$

$$\mathcal{R} = \mathbf{R}_1 + \mathbf{R}_2\mathbf{L} + 2\,(\,r\,\mathbf{e}\mathbf{e}^T - 2\,\mathbf{r}_1\mathbf{e}^T\,) \tag{21}$$

Again, a steepest descent algorithm can be written:

$$\tilde{\mathbf{a}}_1(k+1) = \tilde{\mathbf{a}}_1(k) + 2\mu\,(\,\mathcal{D} - \mathcal{R}_s\,\tilde{\mathbf{a}}_1(k)\,) \tag{22}$$

where $\mathcal{R}_s$ is the symmetric part of matrix $\mathcal{R}$. The instantaneous estimates for $\mathcal{D}$ and $\mathcal{R}_s$ can be derived as previously. Let $\mathbf{u}$ be the following $[(M-1)/2 \times 1]$ vector:

$$\mathbf{u} = [n_{(1)} + n_{(M)}, \ldots, n_{((M-1)/2)} + n_{((M+3)/2)}] \quad (23)$$

then:

$$\hat{p}(k) = n^k_{((M+1)/2)} \{ n^k_{((M+1)/2)}\mathbf{e} - \mathbf{n}_r(k) \} \quad (24)$$

$$\hat{R}_s(k) = \mathbf{n}_{r1}(k)\, \mathbf{u}^T(k) - 2\, n^k_{((M+1)/2)}\mathbf{e}\, \mathbf{n}^T_{r1}(k) +$$

$$+ 2\, n^k_{((M+1)/2)} \{ n^k_{((M+1)/2)}\mathbf{e} - \mathbf{n}_{r1}(k) \}\, \mathbf{e}^T \quad (25)$$

Let $\mathbf{w}(k) = [n^k_{(1)} - n^k_{(M)}, \ldots, n^k_{((M-1)/2)} - n^k_{((M+3)/2)}]^T$ and $v(k) = \mathbf{a}_1(k)\, \mathbf{w}(k)$. After some algebraic manipulation the following unbiased LMS adaptation formula is obtained:

$$\hat{\mathbf{a}}_1(k+1) = \hat{\mathbf{a}}_1(k) + 2\mu\, \{ \varepsilon(k)\, [\hat{\mathbf{n}}_{r1}(k) - n^k_{((M+1)/2)}\mathbf{e}] +$$

$$+ v(k)\, n^k_{((M+1)/2)}\, \mathbf{e} \} \quad (26)$$

3 Convergence properties of the proposed adaptive L-filters

In this section , the convergence in the mean of the location-invariant adaptive LMS L-filter by using the fundamental assumption [1] is proven. The proof of the convergence for the unbiased adaptive LMS L-filter is similar.

Let $\tilde{\mathbf{a}}_o$ denote the vector of the optimal coefficients except the one for the median sample. Then:

$$\mathbf{R}'_s\, \tilde{\mathbf{a}}_o = \mathbf{p}' \quad (27)$$

The optimal filter coefficient for the median sample is given by $a_{(M+1)/2} = 1 - \mathbf{e}^T_{M-1}\, \tilde{\mathbf{a}}_o$. Let $\mathbf{c}(k)$ denote the coefficient error vector:

$$\mathbf{c}(k) = \hat{\mathbf{a}}(k) - \tilde{\mathbf{a}}_o \quad (28)$$

The coefficient error for $a_{(M+1)/2}$ is:

$$c_{(M+1)/2} = -\, \mathbf{e}^T_{M-1}\, \mathbf{c}\,(k) \quad (29)$$

Thus it suffices to prove that E[c(k)] tends to 0 as k tends to ∞. If the LMS algorithm is rewritten in terms of $\mathbf{c}(k)$ and the expected values are taken, then by using the independence of the coefficient vector from the previous ordered sample vectors the following equation is obtained:

$$E[\mathbf{c}(k+1)] = (\mathbf{I} - \mu\, \mathbf{R}'_s)\, E[\mathbf{c}(k)] + \mu\, (\mathbf{p}' - \mathbf{R}'_s\, \tilde{\mathbf{a}}_o) \quad (30)$$

The second term of the right side of (30) equals zero due to (27). Therefore :

$$E[\mathbf{c}(k+1)] = (\, \mathbf{I} - \mu\, \mathbf{R}'_s\,)\, E[\mathbf{c}(k)] \quad (31)$$

Therefore the mean of $\mathbf{c}(k)$ converges to zero as k tends to ∞, when $\mathbf{R}'_s$ is positive definite and the the following inequality is satisfied:

$$0 < \mu < \frac{2}{\lambda_{max}} \quad (32)$$

where λ_{max} is the largest eigenvalue of the matrix $\mathbf{R}'_s$. Since $\mathbf{R}'_s$ is characterized by a large eigenvalue spread, the rate of convergence of the location-invariant LMS adaptive L-filter determined by its minimal eigenvalue.

4 Simulation Examples

The proposed LMS constrained adaptive L-filters have been implemented using C language and have been tested for one dimensional signals for uniform, Gaussian and Laplacian distributions. As a measure for convergence, we shall consider the coefficient estimation error $\Lambda(\mathbf{a},k)$, which is defined by:

$$\Lambda(\mathbf{a},k) = \frac{1}{M} \sum_{j=1}^{M} (\, a_j(k) - a_{j,o}\,)^2 \quad (33)$$

where $a_{j,o}$, j=1,...,M are the optimal coefficients for uniform, Gaussian and Laplacian distribution reported in [5]. It has been shown that the proposed nonlinear adaptive filters can adapt well to a variety of noise probability distributions ranging from the short-tailed ones (e.g. uniform) to the long-tailed ones (e.g. Laplacian).

An example of the performance of the location-invariant LMS L-filter is shown in Figure 2. The input constant signal corrupted by Laplacian white noise having zero mean and variance 2.0 is shown in Figure 2a. The corresponding output of the location-invariant LMS L-filter for M=9 is shown in Figure 2b. The initial L-filter is chosen to be midpoint (i.e. $a_1 = a_9 = 0.5$, $a_2 = \ldots = a_8 = 0$). The choice for the adaptation step is μ=0.003. The coefficient estimation error $\Lambda(\mathbf{a},k)$ is plotted in Figure 2c.

Acknowledgements

This work was supported by the National Scholarship Foundation of Greece and by the Bodosaki Foundation.

References

[1] S. Haykin, Adaptive filter theory, Prentice-Hall, Englewood Cliffs, 1986.

[2] I. Pitas, A.N. Venetsanopoulos, Nonlinear digital filters:Principles and applications, Kluwer Academic, 1990.

[3] I. Pitas, A.N. Venetsanopoulos, "Adaptive filters based on order statistics", IEEE Transactions on ASSP, in press.

[4] F. Palmieri, " A backpropagation algorithm for multilayer hybrid order statistic filters", Proc. of the Int. Conf. on ASSP, Glasgow, England, May 1989.

[5] A. Bovik, T.S. Huang, D.C. Munson, "A generalization of median filtering using linear combinations of order statistics", IEEE Transactions on ASSP, Vol. ASSP-31, No. 6, December 1983.

[6] F.Palmieri, C.G. Boncelet Jr.," A new class of order statistic filters ", Transactions on ASSP, vol. ASSP-37, No. 5, pp. 691-701, May 1989.

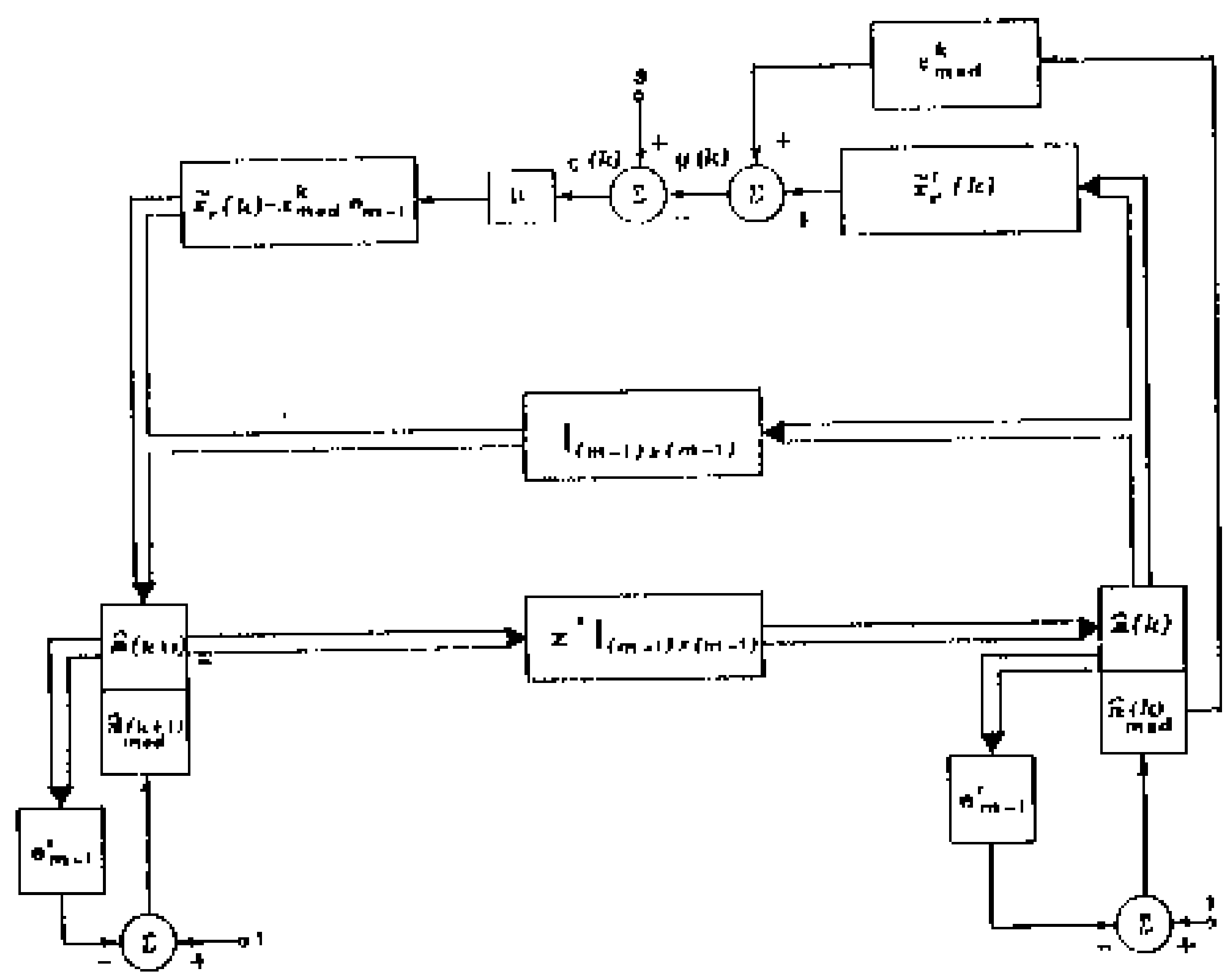
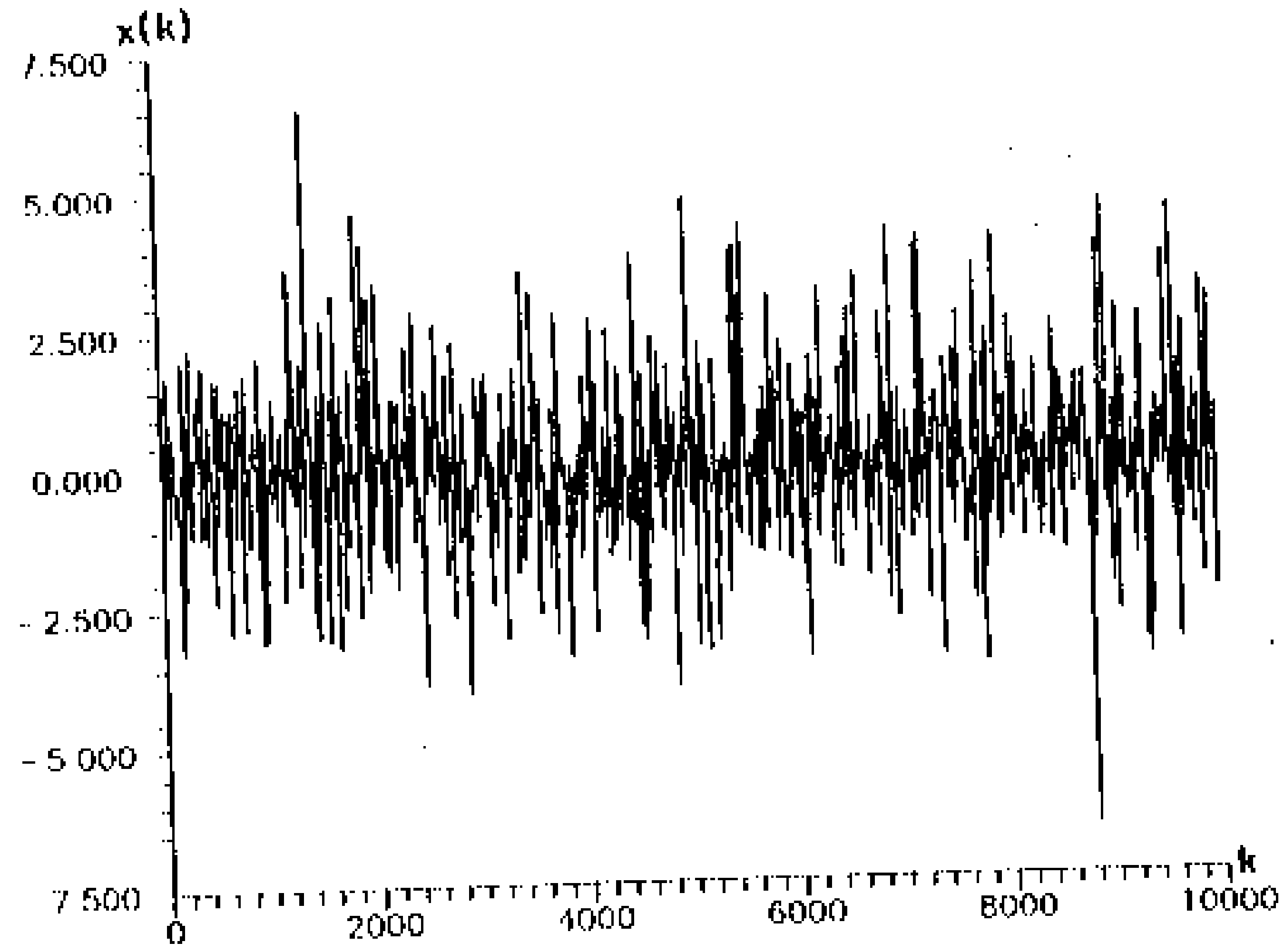
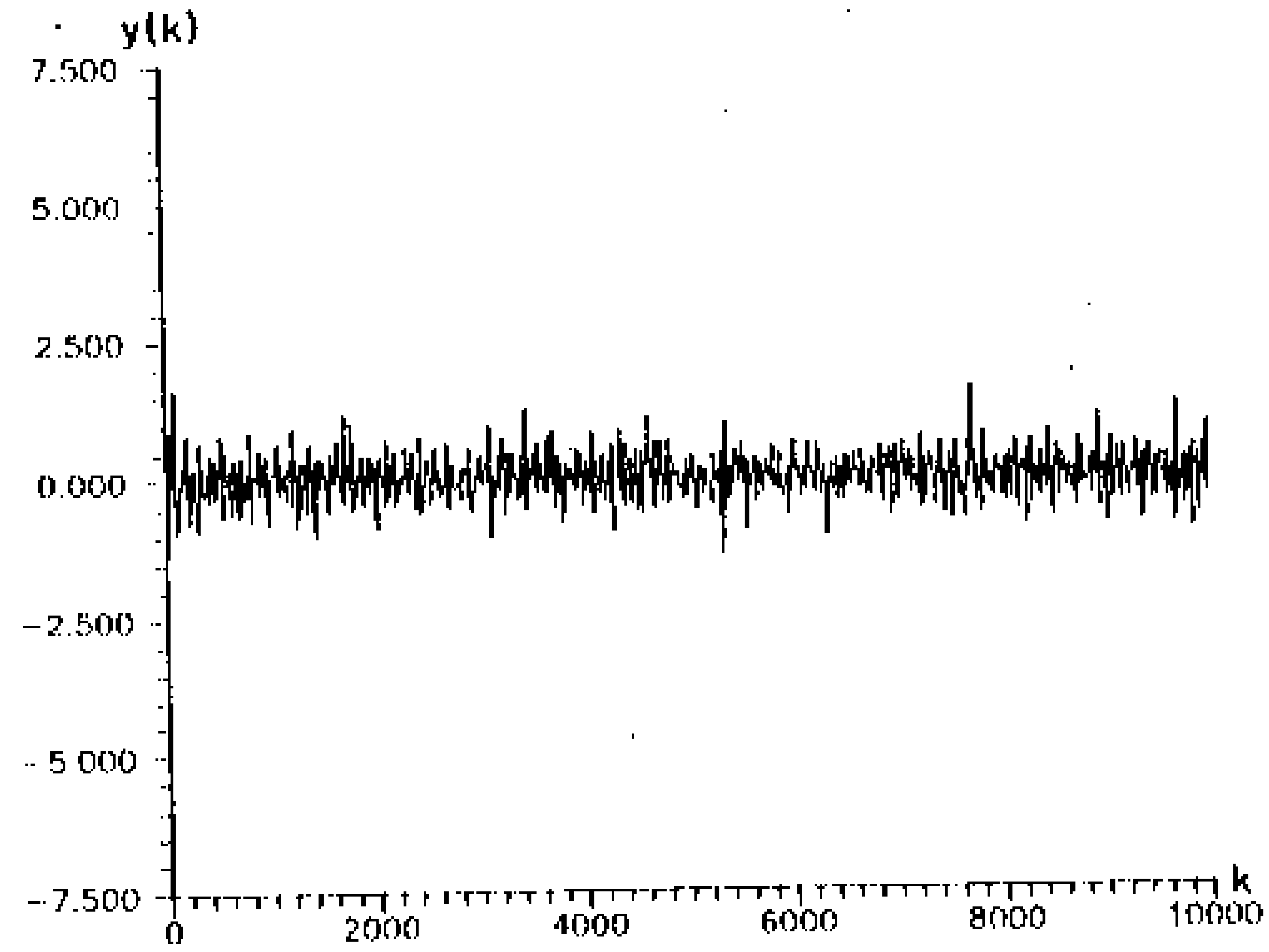


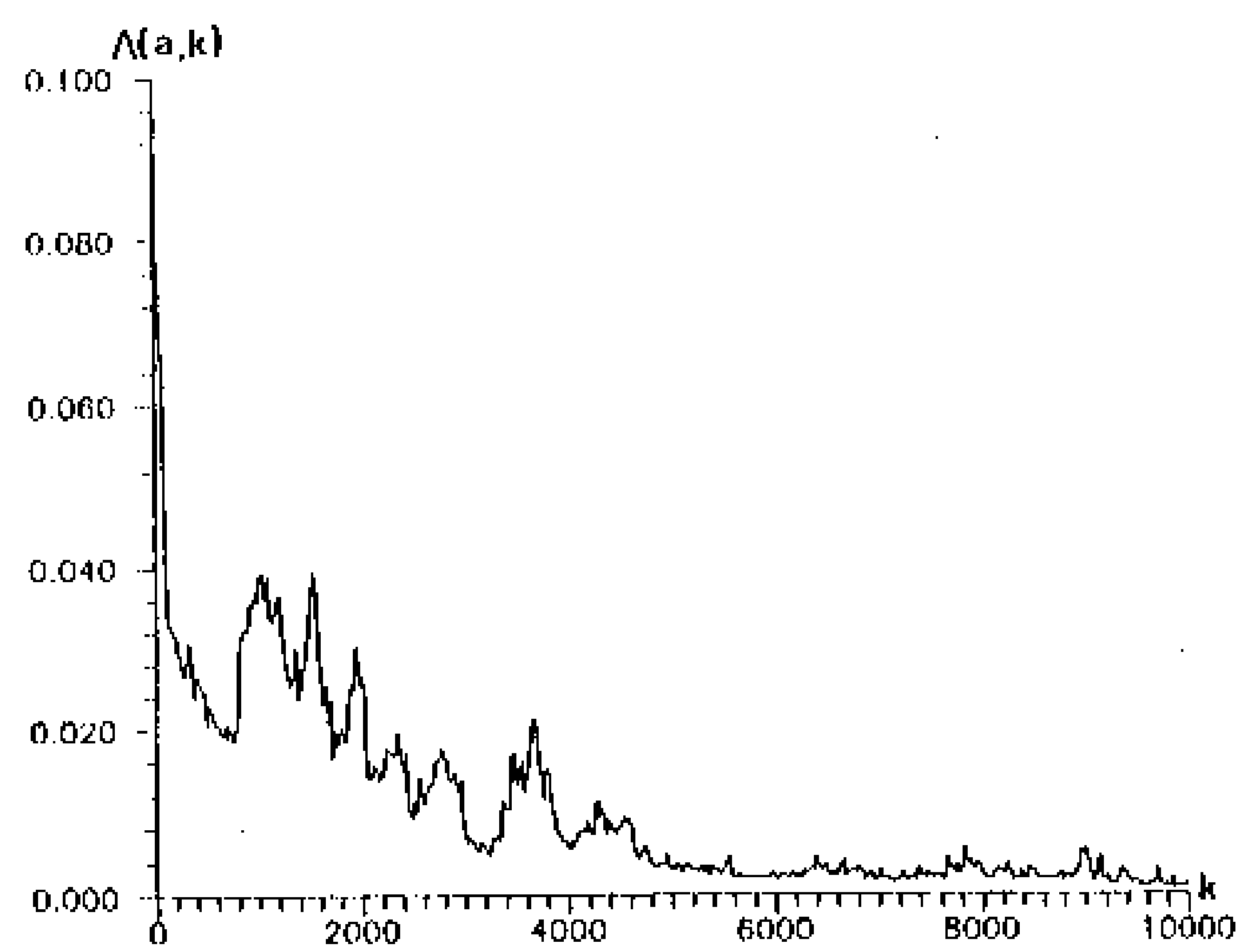
Figure 1: Location-invariant LMS L-filter structure.



(a)



(b)



(c)

Figure 2:(a) Laplacian noise having zero mean and variance 2.0. (b) Filter output by using location-invariant LMS L-filter for M=9, when the initial filter is midpoint. (c) Coefficient estimation error of the location-invariant LMS L-filter for Laplacian noise, when the initial filter is midpoint (M=9).