ADAPTIVE MULTICHANNEL L -FILTERS WITH STRUCTURAL CONSTRAINTS

Constantine Kotropoulos and Maria Gabrani† *and Ioannis Pitas*

Department of Informatics, Aristotle University of Thessaloniki
Box 451, Thessaloniki 540 06, GREECE.
E-mail: {costas, pitas}@zeus.csd.auth.gr
† Department of Electrical and Computer Engineering, Drexel University,
32nd and Chestnut Streets, Philadelphia, PA 19104, U.S.A.
E-mail: maria@cbis.ece.drexel.edu

ABSTRACT

Adaptive multichannel L-filters based on marginal ordering are studied in this paper when structural constraints such as the location-invariance or the unbiasedness are imposed on the filter coefficients. Two novel adaptive algorithms are derived by using Frost's algorithm for minimizing the Mean Squared Error subject to the above-mentioned constraints in the LMS and in the LMS-Newton algorithms. It is demostrated by experiments that the Frost-LMS algorithm has a faster convergence rate than the Frost LMS-Newton algorithm but it yields a higher steady-state MSE than that.

1. INTRODUCTION

Adaptive signal processing has been an active research area for more than two decades. Adaptive filters have been applied in a wide variety of problems including system identification, channel equalization, echo cancellation in telephone channels [1]. All the above-mentioned problems involve one-dimensional (1-D) signals and 1-D linear filters. However, many digital signal processing problems cannot be solved by using linear techniques. Such problems are related to nonlinearities due to noise and/or signal statistics, to system nonlinearities in digital signal acquisition etc. In this case, a multitude of nonlinear techniques has been proved a successful alternative to the linear ones [2].

One of the best known nonlinear filter classes is based on the *order statistics*. It uses the concept of data ordering. There is now a plethora of nonlinear filters based on data ordering. Among them are the L-filters whose output is defined as a linear combination of the order statistics [3]. Recently, increasing attention has been given to nonlinear processing of vector-valued signals [4, 5, 6, 7, 8]. The major difficulty in the definition of multichannel order statistics filters is the lack of an unambiguous and universally accepted definition of ordering for multivariate data [9]. Filters such as those proposed in [6, 7] are based on marginal ordering whereas other filters are based on reduced ordering [4, 5, 8].

The main contribution of this paper is in the design of adaptive multichannel L-filters based on marginal data when structural constraints are imposed on the filter coefficients. In other words, the fidelity criterion to be minimized is the Mean Squared Error (MSE) between the filter output and the desired response subject to a set of constraints on the filter coefficients known as *structural constraints*. Two such constraints are the *location invariance* and the *unbiasedness*. Two novel adaptive algorithms are derived by using the algorithm of Frost [10] for minimizing the MSE subject to the above-mentioned constraints in the LMS and in the LMS-Newton algorithms. By examining the learning curves, it is found that the Frost-LMS algorithm exhibits a faster convergence rate than the Frost LMS-Newton algorithm but to a higher steady-state MSE. The work presented in this paper extends previously reported work [7].

The outline of the paper is as follows. Section 2 describes the problem treated and our motivation for developing constrained adaptive multichannel L-filters. The updating equations for the filter coefficients are derived in Section 3. Finally, experimental results are included in Section 4.

2. PROBLEM STATEMENT

The output of a p-channel L-filter of length N operating on a sequence of p-dimensional vectors $\{\mathbf{x}(k)\}$ for N odd is given by [7]:

$$\mathbf{y}(k) \stackrel{\triangle}{=} \mathbf{T}[\mathbf{x}(k)] = \sum_{i=1}^{p} \mathbf{A}_i \tilde{\mathbf{x}}_i(k) \tag{1}$$

where $\mathbf{A}_i$ is a $(p \times N)$ coefficient matrix. Let $\mathbf{a}_{il}^T$, $l = 1, \ldots, p$ denote the l-th row of matrix $\mathbf{A}_i$ and $\tilde{\mathbf{x}}_i(k) = \left(x_{i(1)}(k), \ldots, x_{i(N)}(k)\right)^T$, $i = 1, \ldots, p$ be the $(N \times 1)$ vector of the order statistics along the i-th channel. Let also $\mathbf{a}_{(i)} = \left(\mathbf{a}_{1i}^T \mid \mathbf{a}_{2i}^T \mid \cdots \mid \mathbf{a}_{pi}^T\right)^T$. Moreover, we define the composite vector $\tilde{\mathbf{X}}(k) = (\tilde{\mathbf{x}}_1^T(k) \mid \tilde{\mathbf{x}}_2^T(k) \mid \ldots \mid \tilde{\mathbf{x}}_p^T(k))^T$.

Frequently, structural constraints are imposed on the filter coefficients. Two such constraints are the *location invariance* and the *unbiasedness*. In the single-channel case, location invariance implies that the sum of filter coefficients must be equal to one. Such a constraint is imposed both to single-channel linear adaptive filters (e.g. the two-dimensional Least Mean Squares (TDLMS) adaptive filters [11]) and to single-channel nonlinear adaptive filters (e.g.

the location-invariant LMS L-filter [12]). In the multichannel case, the optimal nonadaptive location-invariant L-filter has been derived in [7]. Let us recall the definition of the location-invariant multichannel L-filter first. A multichannel marginal L-filter is said to be *location-invariant* [7] if its output is able to track small perturbations of its input. That is, if $\mathbf{x}'(k) = \mathbf{x}(k) + \mathbf{b}$ then:

$$\mathbf{y}'(k) = \mathbf{T}[\mathbf{x}'(k)] = \mathbf{y}(k) + \mathbf{b} \tag{2}$$

where $\mathbf{y}(k) = \mathbf{T}[\mathbf{x}(k)]$. The definition of a location-invariant multichannel L-filter yields the following set of constraints imposed on the filter coefficients:

$$\mathbf{G}^T \mathbf{a}_{(i)} = \mathbf{b}_i \qquad i = 1, \ldots, p \tag{3}$$

where $\mathbf{G}^T$ is a $(p \times pN)$ matrix having the structure:

$$\mathbf{G}^T = \begin{bmatrix} \mathbf{1}^T & \mathbf{0}^T & \cdots & \mathbf{0}^T \\ \vdots & & \ddots & \vdots \\ \mathbf{0}^T & \mathbf{0}^T & \cdots & \mathbf{1}^T \end{bmatrix} \tag{4}$$

with $\mathbf{0}$ being a $(N \times 1)$ vector of zeroes. In (3), $\mathbf{b}_i$ is the i-th basis vector in $\mathcal{R}^p$, i.e., a vector whose elements are zero except the i-th element which equals 1. Another constraint used in practice is the unbiasedness. A multichannel marginal L-filter is said to be an *unbiased* estimator of location if $\mathrm{E}[\mathbf{y}(k)] = \mathrm{E}[\mathbf{s}(k)]$, i.e.,

$$\mathbf{a}_{(i)}^T \, \mathrm{E}[\tilde{\mathbf{X}}(k)] = \mathrm{E}[s_i(k)]. \tag{5}$$

Let us suppose that the observed p-dimensional signal $\{\mathbf{x}(k)\}$ can be expressed as a sum of a p-dimensional noise-free signal $\{\mathbf{s}(k)\}$ and a noise vector sequence $\{\mathbf{n}(k)\}$ of zero mean vector having the same dimensionality, i.e., $\mathbf{x}(k) = \mathbf{s}(k) + \mathbf{n}(k)$. The noise vector components are assumed to be uncorrelated in the general case. In addition, we assume that the noise vectors at different values of index k are independent identically distributed (i.i.d.) and that at each value of index k the signal vector $\mathbf{s}(k)$ and the noise vector $\mathbf{n}(k)$ are uncorrelated. We want to find the multichannel L-filter coefficient matrices $\mathbf{A}_i$, $i = 1, \ldots, p$ which minimize the MSE between the filter output $\mathbf{y}(k)$ and the noise-free signal $\mathbf{s}(k)$ subject to the constraints (3) or (5). Following similar reasoning as in [7], but without invoking the assumption of a constant signal $\mathbf{s}$, it can be shown that the MSE is expressed as:

$$\varepsilon(k) = \sum_{i=1}^{p} \{\mathbf{a}_{(i)}^T \tilde{\mathbf{R}}_p \mathbf{a}_{(i)} - 2\mathbf{a}_{(i)}^T \tilde{\mathbf{q}}_{(i)}\} + \mathrm{E}\left[\mathbf{s}^T(k)\mathbf{s}(k)\right] \tag{6}$$

where $\tilde{\mathbf{R}}_p = \mathrm{E}\left[\tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}^T(k)\right]$ and $\tilde{\mathbf{q}}_{(i)} = \mathrm{E}\left[s_i(k)\tilde{\mathbf{X}}^T(k)\right]$. It can easily be seen that $\tilde{\mathbf{R}}_p$ is a composite matrix that consists of the correlation matrices of the ordered input samples from the same channel (e.g. $\mathbf{R}_{ii} = \mathrm{E}[\tilde{\mathbf{x}}_i(k)\tilde{\mathbf{x}}_i^T(k)]$) as well as from different channels (e.g. $\mathbf{R}_{ij} = \mathrm{E}[\tilde{\mathbf{x}}_i(k)\tilde{\mathbf{x}}_j^T(k)]$, $i \neq j$):

$$\tilde{\mathbf{R}}_p = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} & \cdots & \mathbf{R}_{1p} \\ \vdots & & \ddots & \vdots \\ \mathbf{R}_{1p}^T & \mathbf{R}_{2p}^T & \cdots & \mathbf{R}_{pp} \end{bmatrix}. \tag{7}$$

We can always solve the two constrained minimization problems outlined above provided that we are able to calculate the moments of the order statistics from univariate populations that appear in $\mathbf{R}_{ii}$ as well as the product moments of the order statistics from bivariate populations that appear in $\mathbf{R}_{ij}$, $i \neq j$ and $i = 1, \ldots, p$. This is fairly easy for i.i.d. input variates, i.e., in the case of a constant signal $\mathbf{s}(k) = \mathbf{s}$, as has been demonstrated in [7]. Even for independent, non-identically distributed input variates the framework tends to become very complicated (cf. [7]). The difficulties are increased when the observations $\tilde{\mathbf{X}}(k)$ and the desired signal $\mathbf{s}(k)$ are strongly nonstationary. In order to overcome this obstacle, we shall resort on iterative algorithms for the minimization of $\varepsilon(k)$ in (6) subject to constraints (3) or (5).

3. CONSTRAINED MINIMIZATION OF THE MEAN SQUARED ERROR

In this section, the location-invariant and the unbiased adaptive multichannel L-filter based on marginal ordering are derived. For each constrained adaptive multichannel L-filter Frost's approach [10] is used to minimize iteratively the MSE subject to constraints in the framework either of the LMS algorithm or of the LMSN algorithm.

3.1. Location-invariant adaptive multichannel L-filter

Let us treat the minimization of the MSE (6) subject to (3). The problem under study is formulated as the minimization of the following Lagrangian function:

$$H(\mathbf{a}) = \frac{1}{2}\varepsilon(k) + \mathbf{\Lambda}^T \begin{bmatrix} \mathbf{G}^T \mathbf{a}_{(1)} - \mathbf{b}_1 \\ \vdots \\ \mathbf{G}^T \mathbf{a}_{(p)} - \mathbf{b}_p \end{bmatrix} \tag{8}$$

where $\varepsilon(k)$ is given by (6) and $\mathbf{\Lambda} = \left(\boldsymbol{\lambda}_1^T \mid \cdots \mid \boldsymbol{\lambda}_p^T\right)^T$ is a $(p^2 \times 1)$ vector. By differentiating $H(\mathbf{a})$ with respect to $\mathbf{a}_{(i)}$ and by demanding $\mathbf{a}_{(i)}(k+1)$, $i = 1, \ldots, p$ to satisfy the set of constraints (3) we get:

$$\mathbf{a}_{(i)}(k+1) = \mathbf{P}\left\{\mathbf{a}_{(i)}(k) + \mu\left[\tilde{\mathbf{q}}_{(i)} - \tilde{\mathbf{R}}_p \mathbf{a}_{(i)}(k)\right]\right\} + \mathbf{f}_i. \tag{9}$$

$\mathbf{P}$ is the *projection matrix* of dimensions $(pN \times pN)$ defined by:

$$\mathbf{P} = \left[\mathbf{I} - \mathbf{G}(\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\right] = \left[\mathbf{I} - \frac{1}{N}\mathbf{G}\mathbf{G}^T\right] \tag{10}$$

and $\mathbf{f}_i$ is a $(pN \times 1)$ vector given by:

$$\mathbf{f}_i = \mathbf{G}(\mathbf{G}^T\mathbf{G})^{-1}\mathbf{b}_i = \frac{1}{N}\mathbf{G}\mathbf{b}_i \qquad i = 1, \ldots, p. \tag{11}$$

By using instantaneous estimates for $\tilde{\mathbf{R}}_p$ and $\tilde{\mathbf{q}}_{(i)}$, the LMS location-invariant multichannel L-filter is obtained:

$$\hat{\mathbf{a}}_{(i)}(k+1) = \mathbf{P}\left(\hat{\mathbf{a}}_{(i)}(k) + \mu e_i(k)\tilde{\mathbf{X}}(k)\right) + \mathbf{f}_i \qquad i = 1, \ldots, p. \tag{12}$$

It is evident that by replacing $e_i(k)$ by $-y_i(k)$ the same filter structure can be used for the minimization of the total output power subject to constraints (3). Such an approach can be used when a reference signal is not available. The recursion (12) is initialized by:

$$\mathbf{a}_{(i)}(0) = \mathbf{f}_i \qquad i = 1, \ldots, p. \tag{13}$$

The vast majority of constrained adaptive algorithms relies on the LMS algorithm. To the authors' knowledge no attempt has been made to design constrained adaptive filters based on other adaptive algorithms, such as the Recursive Least Squares (RLS) or the LMS Newton (LMSN) algorithm. The case is much simpler for the LMSN [13] than for the RLS algorithm, because LMSN shares the same framework with LMS. It is well known that LMSN minimizes the cost function (8) as well. It can be shown that the optimal solution to the minimization of the cost function (8) is given by:

$$\mathbf{a}^*_{(i)}(k) = \mathbf{a}_{(i)}(k) - \tilde{\mathbf{R}}_p^{-1} \frac{\partial H(\mathbf{a}(k))}{\partial \mathbf{a}_{(i)}(k)}. \tag{14}$$

The steepest descent solution is obtained from (14) by adding an additional step-size parameter μ:

$$\mathbf{a}_{(i)}(k+1) = \mathbf{a}_{(i)}(k) - \mu \tilde{\mathbf{R}}_p^{-1} \frac{\partial H(\mathbf{a}(k))}{\partial \mathbf{a}_{(i)}(k)}. \tag{15}$$

By substituting $\tilde{\mathbf{R}}_p^{-1}$ with the estimate

$$\hat{\mathbf{R}}_p^{-1}(k) = \frac{1}{1-\zeta} \Big\{ \hat{\mathbf{R}}_p^{-1}(k-1) - \frac{\hat{\mathbf{R}}_p^{-1}(k-1)\tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}^T(k)\hat{\mathbf{R}}_p^{-1}(k-1)}{\left(\frac{1-\zeta}{\zeta}\right)\tilde{\mathbf{X}}^T(k)\hat{\mathbf{R}}_p^{-1}(k-1)\tilde{\mathbf{X}}(k)} \Big\} \tag{16}$$

and by using instantaneous estimates for the expected values involved in the gradient of $H(\mathbf{a}(k))$ with respect to $\mathbf{a}_{(i)}(k)$, the following recursions result:

$$\begin{aligned} \hat{\mathbf{a}}_{(i)}(k+1) &= \Big\{ \mathbf{I} - \hat{\mathbf{R}}_p^{-1}(k)\mathbf{G}\left[\mathbf{G}^T\hat{\mathbf{R}}_p^{-1}(k)\mathbf{G}\right]^{-1}\mathbf{G}^T \Big\} \\ &\cdot \left[\hat{\mathbf{a}}_{(i)}(k) + \mu \hat{\mathbf{R}}_p^{-1}(k) e_i(k)\tilde{\mathbf{X}}(k)\right] \\ &+ \hat{\mathbf{R}}_p^{-1}(k)\mathbf{G}\left[\mathbf{G}^T\hat{\mathbf{R}}_p^{-1}(k)\mathbf{G}\right]^{-1}\mathbf{b}_i. \end{aligned} \tag{17}$$

The comparison between (17) and (11) reveals that the structure of the LMSN location-invariant multichannel L-filter is the same with that of the LMS location-invariant one but with a time-varying matrix $\mathbf{P}$ and a time-varying vector $\mathbf{f}_i$. The new matrix $\mathbf{P}'(k)$ and vector $\mathbf{f}'_i(k)$, $i = 1, \ldots, p$ are now given by:

$$\mathbf{P}'(k) = \mathbf{I} - \hat{\mathbf{R}}_p^{-1}(k)\mathbf{G}\left[\mathbf{G}^T\hat{\mathbf{R}}_p^{-1}(k)\mathbf{G}\right]^{-1}\mathbf{G}^T \tag{18}$$

$$\mathbf{f}'_i(k) = \hat{\mathbf{R}}_p^{-1}(k)\mathbf{G}\left[\mathbf{G}^T\hat{\mathbf{R}}_p^{-1}(k)\mathbf{G}\right]^{-1}\mathbf{b}_i \tag{19}$$

The updating equations (17) are also initialized by (13).

3.2. Unbiased adaptive multichannel L -filter

Let $\tilde{\mathbf{m}}_p = \mathrm{E}[\tilde{\mathbf{X}}(k)]$ and $\overline{s_i} = \mathrm{E}[s_i(k)]$. Eq. (5) can be rewritten as:

$$\mathbf{a}^T_{(i)}\,\tilde{\mathbf{m}}_p = \overline{s_i}. \tag{20}$$

We shall assume that $\tilde{\mathbf{m}}_p$ has already been estimated and it is known. For example, one may use the following recursion:

$$\hat{\mathbf{m}}_p(k) = \hat{\mathbf{m}}_p(k-1) + \frac{1}{k}\left(\tilde{\mathbf{X}}(k) - \hat{\mathbf{m}}_p(k-1)\right) \tag{21}$$

with $\hat{\mathbf{m}}_p(0) = \mathbf{0}$. The minimization of the MSE (6) subject to (20) is formulated as the minimization of the following Lagrangian function:

$$H(\mathbf{a}) = \frac{1}{2}\varepsilon(k) + \sum_{i=1}^{p} \lambda_i \left(\mathbf{a}^T_{(i)}\tilde{\mathbf{m}}_p - \overline{s_i}\right). \tag{22}$$

By differentiating $H(\mathbf{a})$ with respect to $\mathbf{a}_{(i)}$ and by demanding $\mathbf{a}_{(i)}(k+1)$ to satisfy the set of constraints (20) we obtain:

$$\begin{aligned} \mathbf{a}_{(i)}(k+1) &= \mathbf{a}_{(i)}(k) - \mu\left(\mathbf{I} - \frac{\tilde{\mathbf{m}}_p\tilde{\mathbf{m}}_p^T}{\tilde{\mathbf{m}}_p^T\tilde{\mathbf{m}}_p}\right)\Big(\tilde{\mathbf{R}}_p\mathbf{a}_{(i)}(k) - \\ &- \mathbf{q}_{(i)}\Big) + \frac{\left(\overline{s_i} - \tilde{\mathbf{m}}_p^T\mathbf{a}_{(i)}(k)\right)}{\tilde{\mathbf{m}}_p^T\tilde{\mathbf{m}}_p}\tilde{\mathbf{m}}_p. \end{aligned} \tag{23}$$

By rearranging the terms in (23) we get:

$$\begin{aligned} \mathbf{a}_{(i)}(k+1) &= \mathbf{P}'' \left\{ \left(\mathbf{a}_{(i)}(k) + \mu(\mathbf{q}_{(i)} - \tilde{\mathbf{R}}_p\mathbf{a}_{(i)}(k)\right) \right\} \\ &+ \mathbf{f}''_i \quad i = 1, \ldots, p \end{aligned} \tag{24}$$

with

$$\mathbf{P}'' = \left(\mathbf{I} - \frac{\tilde{\mathbf{m}}_p\tilde{\mathbf{m}}_p^T}{\tilde{\mathbf{m}}_p^T\tilde{\mathbf{m}}_p}\right) \tag{25}$$

$$\mathbf{f}''_i = \frac{\overline{s_i}}{\tilde{\mathbf{m}}_p^T\tilde{\mathbf{m}}_p}\tilde{\mathbf{m}}_p. \tag{26}$$

The algorithm is initialized by:

$$\mathbf{a}_{(i)}(0) = \mathbf{f}''_i. \tag{27}$$

If instantaneous estimates for $\tilde{\mathbf{R}}_p$ and $\mathbf{q}_{(i)}$ are used, we obtain the LMS unbiased multichannel L-filter:

$$\begin{aligned} \hat{\mathbf{a}}_{(i)}(k+1) &= \mathbf{P}''(k)\left\{\hat{\mathbf{a}}_{(i)}(k) + \mu e_i(k)\tilde{\mathbf{X}}(k)\right\} \\ &+ \mathbf{f}''_i(k) \quad i = 1, \ldots, p \end{aligned} \tag{28}$$

Let us also consider the minimization of the MSE (6) subject to (20) within the framework of LMS-Newton algorithm. Following similar reasoning it can easily be proved that:

$$\begin{aligned} \hat{\mathbf{a}}_{(i)}(k+1) &= \hat{\mathbf{P}}'''(k)\left\{\hat{\mathbf{a}}_{(i)}(k) + \mu\hat{\mathbf{R}}_p^{-1}(k)e_i(k)\tilde{\mathbf{X}}(k)\right\} \\ &+ \hat{\mathbf{f}}'''_i(k) \quad i = 1, \ldots, p \end{aligned} \tag{29}$$

with

$$\hat{\mathbf{P}}'''(k) = \left(\mathbf{I} - \frac{\hat{\mathbf{R}}_p^{-1}(k)\tilde{\mathbf{m}}_p(k)\tilde{\mathbf{m}}_p^T(k)}{\tilde{\mathbf{m}}_p^T(k)\hat{\mathbf{R}}_p^{-1}(k)\tilde{\mathbf{m}}_p(k)}\right) \tag{30}$$

$$\hat{\mathbf{f}}'''_i(k) = \frac{\overline{s_i}\hat{\mathbf{R}}_p^{-1}(k)\tilde{\mathbf{m}}_p(k)}{\tilde{\mathbf{m}}_p^T(k)\hat{\mathbf{R}}_p^{-1}(k)\tilde{\mathbf{m}}_p(k)}. \tag{31}$$

The updating equations (29) can be initialized by:

$$\mathbf{a}_{(i)}(0) = \mathbf{f}''_i \tag{32}$$

as well. Eqs. (29)–(32) define the LMSN unbiased multichannel L-filter.

4. EXPERIMENTAL RESULTS

A set of experiments is presented in order to assess the performance of the location-invariant adaptive multichannel L-filters that we have discussed so far.

A two-channel 1-D signal $\mathbf{s}(k) = \mathbf{s}$ corrupted by additive white bivariate contaminated Gaussian noise is treated, because for such a signal, the optimal multichannel L-filter coefficients have been derived in [7]. Let $\mathcal{N}(\xi_1, \xi_2\,; \sigma_1, \sigma_2\,; r)$ denote a joint bivariate Gaussian distribution where the parameter ξ_i and σ_i, $i = 1, 2$ is the expected value and the standard deviation of each component respectively and r is the correlation coefficient. A vector valued signal $\mathbf{s} = (1.0, 2.0)^T$ corrupted by additive white bivariate noise $\mathbf{n}(k)$ with probability density function (pdf) given by:

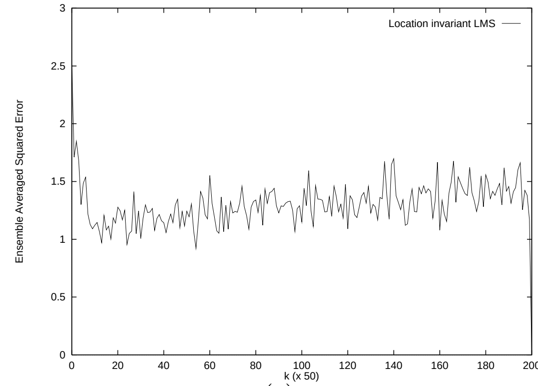
$$(1-\varrho)\mathcal{N}(0,0\,;1,3\,;0.5) + \varrho\mathcal{N}(0,0\,;3,9\,;0.7)$$

for $\varrho = 0.1$ has been used as a test signal as in [7]. The noise reduction index (NR) defined as the ratio of the output noise power to the input noise power, i.e.:

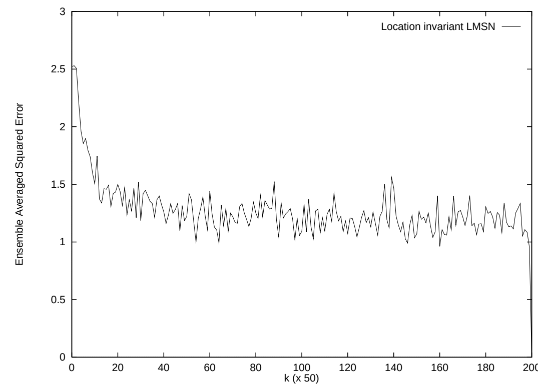
$$\mathrm{NR} = 10 \log \frac{\sum_k (\mathbf{y}(k) - \mathbf{s}(k))^T (\mathbf{y}(k) - \mathbf{s}(k))}{\sum_k (\mathbf{x}(k) - \mathbf{s}(k))^T (\mathbf{x}(k) - \mathbf{s}(k))}. \quad (33)$$

is measured and is compared to the one achieved by the nonadaptive location-invariant multichannel L-filter.

An approximation of the ensemble-averaged learning curve for the location-invariant multichannel L-filters under study has been obtained following the procedure described in [1]. That is, a sequence of 10000 samples of $\{\mathbf{x}(k)\}$ has been created and the squared norm of the estimation error $\|\mathbf{e}(k)\|^2 = \|\mathbf{s}(k) - \mathbf{y}(k)\|^2$ has been computed. The experiment has been repeated 200 times, each time using an independent realization of the process $\{\mathbf{n}(k)\}$. The averaged squared norm of the estimation error is then determined by computing the ensemble average of $\|\mathbf{e}(k)\|^2$ over the 200 independent trials of the experiment. The learning curves of the LMS and LMSN location-invariant multichannel L-filters are given in Figure 1a and 1b, respectively. The filter length N has been 9 in all cases. The recursions have been initialized by using (13). In the plots of Figure 1 points every 50 time instants have been used. In the location-invariant LMS multichannel L-filter, μ has been equal to 5×10^{-5}. For the location-invariant LMSN multichannel L-filter, μ, ζ have been set to 0.001 and 0.001, respectively. The recursion for the computation of the inverse matrix starts with $\hat{\mathbf{R}}_p^{-1}(0) = \delta^{-1}\mathbf{I}$ with $\delta = 0.01$. It is seen that LMS exhibits a faster convergence rate but to a higher steady state MSE, as is manifested in Table 1 where the NR achieved by both the adaptive algorithms is tabulated. In the same table, the NR achieved by the nonadaptive design [7] is included for comparison purposes. The NR achieved by the marginal median is included for the same purposes as well. The LMSN location-invariant multichannel L-filter outperforms the nonadaptive one by 1.4 dB. This is attributed to the arithmetic errors that occur in the estimation of the moments of the marginal order statistics (i.e., numerical integration or discretization) employed in [7]. The very large eigenvalue spread that is inherent to the correlation matrix of the order statistics magnifies these errors. As a consequence, the filter coefficients are seriously effected.



(a)



(b)

Figure 1: Learning curves of the (a) LMS adaptive location-invariant multichannel L-filter; (b) LMSN adaptive location-invariant multichannel L-filter.

The learning curves of the LMS and LMSN unbiased multichannel L-filters have been computed as well. They are shown in Figure 2a and 2b, respectively. A variable step-size $\mu(k) = 0.1 \times (\tilde{\mathbf{X}}^T\tilde{\mathbf{X}})^{-1}$ has been used. It is seen that LMS converges faster than LMSN algorithm but to a higher steady-state MSE in this case, too. Table 2 summarizes the NR achieved by both the adaptive algorithms. For comparison purposes the NR achieved by the nonadaptive unbiased multichannel L-filter is also tabulated.

5. REFERENCES

[1] S. Haykin, *Adaptive Filter Theory.* Englewood Cliffs, N.J.: Prentice Hall, 1986.

[2] I. Pitas, and A.N. Venetsanopoulos, *Nonlinear Digital Filters: Principles and Applications.* Norwell, MA: Kluwer Academic Publ., 1990.

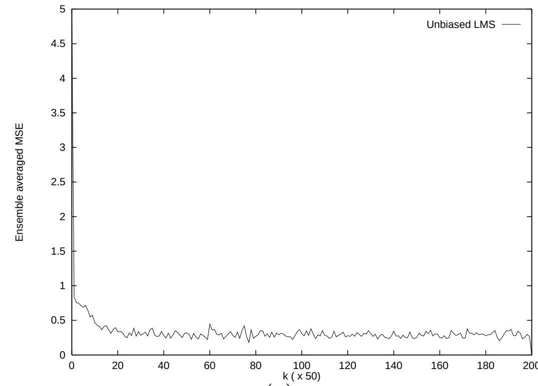
[3] A.C. Bovik, T.S. Huang, and D.C. Munson, Jr., "A generalization of median filtering using linear combinations of order statistics," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. ASSP-31, no. 6, pp. 1342–1350, December 1983.

Table 1: Noise reduction (in dB) achieved by the location-invariant multichannel L-filters for the bivariate contaminated Gaussian noise model (Filter length $N = 9$).

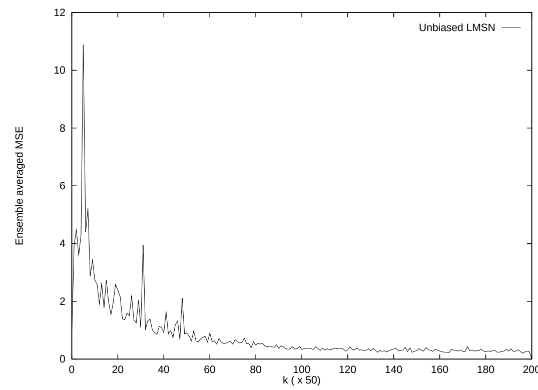
Filter	NR
LMS location-invariant	-11.098
LMSN location-invariant	-12.195
nonadaptive location-invariant	-10.997
marginal median	-9.8209

Table 2: Noise reduction (in dB) achieved by the unbiased multichannel L-filters for the bivariate contaminated Gaussian noise model (Filter length $N = 9$).

Filter	NR
LMS unbiased	-18.144
LMSN unbiased	-18.345
nonadaptive unbiased	-18.346



(a)



(b)

Figure 2: Learning curves of the (a) LMS adaptive unbiased multichannel L-filter; (b) LMSN adaptive unbiased multichannel L-filter.

[4] J. Astola, P. Haavisto, and Y. Neuvo, “Vector median filters,” *Proceedings of the IEEE*, vol. 78, no. 4, pp. 678–689, April 1990.

[5] R.C. Hardie, and G.R. Arce. “Ranking in R^p and its use in multivariate image estimation,” *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 1, no. 2, pp. 197–209, June 1991.

[6] I. Pitas, and P. Tsakalides, “Multivariate ordering in color image restoration,” *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 1, no. 3, pp. 247–259, September 1991.

[7] C. Kotropoulos, and I. Pitas, “Multichannel L-filters based on marginal data ordering,” *IEEE Trans. on Signal Processing*, vol. 42, no. 10, pp. 2581–2595, October 1994.

[8] K.N. Plataniotis, D. Androutsos, and A.N. Venetsanopoulos, “Multichannel filters for image processing,” *Signal Processing: Image Communication*, vol. 9, pp. 143–158, 1997.

[9] V. Barnett, “The ordering of multivariate data,” *J. R. Statist. Soc. A*, vol. 139, Part 3, pp. 318–354, 1976.

[10] O.L. Frost, III, “An algorithm for linearly constrained adaptive array processing,” *Proceedings of the IEEE*, vol. 60, pp. 926–935, 1972.

[11] M.M. Hadhoud, and D. W. Thomas, “The two-dimensional adaptive LMS (TDLMS) algorithm,” *IEEE Trans. on Circuits and Systems*, vol. 35, no. 5, pp. 485–494, May 1988.

[12] C. Kotropoulos, and I. Pitas, “Adaptive LMS L-filters for noise suppression in images,” *IEEE Trans. on Image Processing*, vol. 5, no. 12, December 1996, to appear.

[13] G.V. Moustakides, and S. Theodoridis, “Fast Newton Transversal Filters - A new class of adaptive estimation algorithms,” *IEEE Trans. on Signal Processing*, vol. 39, no. 10, pp. 2184–2193, October 1991.