Multichannel L-filters based on M-ordering

C. Kotropoulos* I. Pitas
Department of Electrical Engineering, University of Thessaloniki
Thessaloniki 540 06, GREECE (fax: +30-31-274868)

Abstract

The extension of single-channel nonlinear filters whose output is a linear combination of the order statistics of the input samples to the multichannel case is presented in this paper. The subordering principle of marginal ordering (M-ordering) is used for multivariate data ordering. A unified framework for a discrete calculation of the moments of the bivariate order statistics required for the design of the multichannel marginal L-filters is outlined. The derivation of a bivariate distribution, namely the Laplacian (bi-exponential) distribution which belongs to Morgenstern's family is discussed. It is shown by simulations that the proposed multichannel L-filters perform better than other multichannel nonlinear filters such as the marginal median and the vector median proposed elsewhere as well as their single-channel counterparts.

1 Introduction

Multichannel one-dimensional and two dimensional signals appear frequently in practice, for example in the cases involving multiple sources and receivers, as in geophysics, underwater acoustics, multiple-antenna transmission systems as well as in the processing of color images and sequences of images. A multichannel signal is defined as a vector of components called channels which are generally correlated and characterized by their joint probability density function (pdf). If each signal component is processed separately, this correlation is not utilized. Although transformation techniques such as the Karhunen-Loeve transformation can be used first to decorrelate the signal components in order to apply single-channel signal processing techniques afterwards, a more natural way is to apply multichannel signal processing techniques.

Single-channel nonlinear filtering techniques have exhibited a tremendous growth in the past decade as alternatives of linear filtering in problems that cannot be efficiently solved by using linear techniques, e.g. in the case of non-Gaussian or signal-dependent noise filtering [5]. A class of nonlinear filters that has found extensive applications in digital signal and image processing are the L-filters (sometimes also called order statistic filters) whose output is defined as a linear combination of the order statistics of the input sequence [8, 12]. Recently, increasing attention has been given to nonlinear processing of vector-valued signals [10, 11, 13, 14].

The main contribution of the paper is the design of multichannel L-filters based on marginal ordering (M-ordering) using the Mean-Squared-Error (MSE) as fidelity criterion. M-ordering implies independent data ordering in each channel. We assume that a multichannel constant signal is corrupted by additive white multivariate noise which generally exhibits correlation between different channels. The unconstrained minimization of the MSE is treated first. Structural constraints such as unbiasedness and location-invariance are also incorporated in the minimization procedure. The unconstrained minimization is shown that it leads to a global minimum. The design procedure involves moments of the order statistics of input samples derived from the same channel as well as from different channels. The theoretical framework required for the computation of the abovementioned moments is outlined and a discrete algorithm for their computation is derived based on vector quantization. In order to test the performance of the designed multichannel marginal L-filters, long-tailed multivariate distributions are required. The derivation and design of such a distribution, namely, the Laplacian (bi-exponential) distribution which belongs to Morgenstern's family in the two-channel case is discussed. The noise reduction capability of the designed multichannel nonlinear filters is examined for bivariate distributions ranging from the short-tailed to the long-tailed ones. The following bivariate pdfs are considered: uniform, joint Gaussian, contaminated Gaussian and Laplacian distributions. It is shown by simulations that the proposed multichannel L-filters perform better than other multichannel nonlinear filters such as the marginal median, the vector median proposed in [10] [13] as well as their single-

*Supported by a scholarship from the State Scholarship Foundation of Greece and the Bodosaki Foundation

channel counterparts.

The work presented in this paper extends previously reported work [10]–[13]. The outline of the paper is as follows. The basic concepts of multivariate data ordering are reviewed in Section 2. The design of multichannel marginal L-filters is also described in this section. The calculation of the moments of multivariate order statistics is treated in Section 3. The derivation of the bivariate Laplacian distribution is outlined in Section 4. Simulation examples are included and conclusions are drawn in Section 5.

2 Multichannel L-filters based on marginal data ordering

The notion of data ordering cannot be extended in a straightforward manner from the univariate case to the multivariate one. An excellent treatment of the multivariate data ordering and the outliers in multivariate data can be found in [7]. There are several ways to order multivariate data, but none of them is unambiguous nor universally accepted. Specifically, the following four so-called sub-ordering principles are discussed in [7]: marginal ordering, partial ordering, conditional ordering and reduced ordering. In the sequel, we shall confine ourselves to the definition of M-ordering.

Let $\mathbf{x}_1, \ldots, \mathbf{x}_N$ be a random sample of N observations of a p-dimensional random variable $\mathbf{X}$ where $\mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{ip})^T$. The M-ordering scheme orders each of the vector components independently yielding:

$$x_{j(1)} \leq x_{j(2)} \leq \cdots \leq x_{j(N)} \quad j = 1, \ldots, p \qquad (1)$$

i.e., the vector-valued observations are ordered along each of the p-dimensions (or channels) independently. By definition, the $(i_1, i_2, \ldots, i_p)$-marginal order statistic is the following $(p \times 1)$ vector:

$$\mathbf{x}_{(i_1, i_2, \ldots, i_p)} \stackrel{\text{def}}{=} (x_{1(i_1)}, x_{2(i_2)}, \ldots, x_{p(i_p)})^T \qquad (2)$$

where $i_j, \ 1 \leq i_j \leq N, \ j = 1, \ldots, p$.

As mentioned before, L-filters have been used extensively as estimators of location in the single-channel case. The following definition for the multichannel L-filter based on marginal ordering has been given in [11]: The output of a p-channel marginal L-filter of length N, $\mathbf{y}(k) = \mathbf{T}[\mathbf{x}(k)]$, operating on a sequence of p-dimensional vectors $\{\mathbf{x}(k)\}$ for N odd is given by:

$$\mathbf{y}(k) \stackrel{\text{def}}{=} \sum_{i_1=1}^{N} \cdots \sum_{i_p=1}^{N} A_{i_1, \ldots, i_p} \mathbf{x}_{(i_1, \ldots, i_p)}(k) \qquad (3)$$

where $A_{i_1, \ldots, i_p}$ are $(p \times p)$ coefficient matrices and the $(i_1, i_2, \ldots, i_p)$-marginal order statistics have been formed by ranking the components of the $(p \times 1)$ vector-valued observations $\mathbf{x}(k - \nu), \ldots, \mathbf{x}(k), \ldots, \mathbf{x}(k + \nu)$ independently along each channel. By rearranging the terms appearing in sum (3), it can be shown that definition (3) is equivalent to :

$$\mathbf{y}(k) = \sum_{j=1}^{p} \mathbf{A}_j \tilde{\mathbf{x}}_j(k) \qquad (4)$$

where $\mathbf{A}_j$ are appropriate $(p \times N)$ coefficient matrices and $\tilde{\mathbf{x}}_j(k) = (x_{j(1)}(k), \ldots, x_{j(N)}(k))^T$ are the $(N \times 1)$ vectors of the order statistics along each channel.

Let us suppose that the observed p-dimensional signal $\{\mathbf{x}(k)\}$ can be expressed as the sum of a known p-dimensional constant signal $\mathbf{s}$ and a noise vector sequence $\{\mathbf{n}(k)\}$ of zero-mean vector having the same dimensionality, i.e., $\mathbf{x}(k) = \mathbf{s} + \mathbf{n}(k)$. The noise vector $\mathbf{n}(k) = (n_1(k), \ldots, n_p(k))^T$ is a p-dimensional vector of random variables characterized by the joint pdf of its components which are assumed to be correlated in the general case. In addition, we assume that the noise vectors at different time-instants are independent, identically distributed (i.i.d.) and that at every time-instant the signal $\mathbf{s}$ and the noise vector $\mathbf{n}(k)$ are uncorrelated.

We shall design the p-channel marginal L-filter which operates on the p-dimensional observed signal $\{\mathbf{x}(k)\}$ and is the optimal estimator of $\mathbf{s}$ by using the MSE between $\mathbf{s}$ and the output of the p-channel marginal L-filter as fidelity criterion.

Let $\mathbf{a}_{jl}^T, \ l = 1, \ldots, p$ be $(1 \times N)$ row vectors corresponding to the rows of matrix $\mathbf{A}_j$. Let $\mathbf{R}_{ji} = \mathrm{E}[\tilde{\mathbf{x}}_j \tilde{\mathbf{x}}_i^T]$ denote the correlation matrix of the ordered input samples in channels j and i. For $j = i$, $\mathbf{R}_{ii} \ \ i = 1, \ldots, p$ consists of moments of the order statistics from a univariate population. For $j \neq i$, the elements of $\mathbf{R}_{ji}$ are moments of the order statistics from a bivariate population. Let also $\underline{\mu}_j = (\mathrm{E}[x_{j(1)}], \mathrm{E}[x_{j(2)}], \ldots, \mathrm{E}[x_{j(N)}])^T$ denote the mean vector of the order statistics in channel j. The MSE between $\mathbf{s}$ and $\mathbf{y}(k)$ is given by:

$$\varepsilon = \sum_{i=1}^{p} \mathbf{a}_{(i)}^T \hat{\mathbf{R}}_p \mathbf{a}_{(i)} - 2\mathbf{s}^T \begin{bmatrix} \mathbf{a}_{(1)}^T \\ \vdots \\ \mathbf{a}_{(p)}^T \end{bmatrix} \hat{\underline{\mu}}_p + \mathbf{s}^T \mathbf{s} \qquad (5)$$

where

$$\mathbf{a}_{(i)} = \begin{pmatrix} \mathbf{a}_{1i}^T & \mathbf{a}_{2i}^T & \cdots & \mathbf{a}_{pi}^T \end{pmatrix}^T \qquad (6)$$

$$\hat{\mathbf{R}}_p = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} & \cdots & \mathbf{R}_{1p} \\ \mathbf{R}_{12}^T & \mathbf{R}_{22} & \cdots & \mathbf{R}_{2p} \\ \vdots & & \ddots & \vdots \\ \mathbf{R}_{1p}^T & \mathbf{R}_{2p}^T & \cdots & \mathbf{R}_{pp} \end{bmatrix} \qquad (7)$$

$$\hat{\underline{\mu}}_p = \begin{pmatrix} \underline{\mu}_1^T & \underline{\mu}_2^T & \cdots & \underline{\mu}_p^T \end{pmatrix}^T \qquad (8)$$

In the sequel we shall treat first the unconstrained minimization of the MSE and then we shall impose constraints on the output of the multichannel marginal L-filter.

Minimizing (5) over $\mathbf{a}_{ji}$, $j, i = 1, \ldots, p$ is a quadratic optimization problem which has a unique solution provided that the symmetric matrix $\hat{\mathbf{R}}_p$ is positive definite. Only the diagonal submatrices of $\hat{\mathbf{R}}_p$ are by definition positive semidefinite [3]. We shall assume that $\hat{\mathbf{R}}_p$ is indeed positive definite. Such an assumption has been verified in all simulations performed in Section 5.

Equating the derivatives of ε with respect to $\mathbf{a}_{ji}$ with zero, i.e., $\frac{\partial \varepsilon}{\partial \mathbf{a}_{ji}} = \mathbf{0}$, the following p sets of equations result:

$$\hat{\mathbf{R}}_p \mathbf{a}^*_{(m)} = s_m \underline{\hat{\mu}}_p \quad m = 1, \ldots, p \tag{9}$$

which yield the optimal p-channel marginal L-filter coefficients:

$$\begin{aligned} \mathbf{a}^*_{(1)} &= s_1 \hat{\mathbf{R}}_p^{-1} \underline{\hat{\mu}}_p \\ \mathbf{a}^*_{(m)} &= \frac{s_m}{s_1} \mathbf{a}^*_{(1)} \quad m = 2, \ldots, p \end{aligned} \tag{10}$$

The MMSE associated with the optimal coefficients (10) is:

$$\varepsilon_{\min} = (1 - \Delta)\mathbf{s}^T\mathbf{s} \; ; \; \Delta = \underline{\hat{\mu}}_p^T \hat{\mathbf{R}}_p^{-1} \underline{\hat{\mu}}_p \tag{11}$$

The fact that ε is always nonnegative implies that $\varepsilon_{\min} \geq 0$. Therefore, $0 \leq \Delta \leq 1$.

In (10), the optimal coefficients depend on the knowledge of the constant signal $\mathbf{s}$. In addition, the joint probability density function of the components of the input vector-valued signal $\mathbf{x}(k)$ must be known in order to calculate $\hat{\mathbf{R}}_p$ and $\underline{\hat{\mu}}_p$ as is analyzed in Section 3. In many practical applications, the constant signal $\mathbf{s}$ is unknown. Therefore, $\mathbf{s}$ must be estimated at every time instant k from the past L-filter outputs $\mathbf{y}(l)$, $l = k-1, k-2, \ldots$. Such an estimate $\hat{\mathbf{s}}(k)$ of $\mathbf{s}$ at time instant k is described in Section 5.

In the univariate case, L-filters are designed by imposing local structural constraints on the output of the L-filter. Two types of constraints have been incorporated in the design of single-channel L-filters [8, 12]: unbiasedness and location-invariance. A multichannel marginal L-filter is said to be unbiased multichannel estimator of location, if $\mathrm{E}[\mathbf{y}(k)] = \mathbf{s}$ holds, or equivalently:

$$\mathbf{a}^T_{(i)} \underline{\hat{\mu}}_p = s_i \quad i = 1, \ldots, p \tag{12}$$

Under the set of constraints (12), the MSE given by (5) is rewritten as:

$$\varepsilon_{\text{unb}} = \sum_{i=1}^{p} \mathbf{a}^T_{(i)} \hat{\mathbf{R}}_p \mathbf{a}_{(i)} - \mathbf{s}^T\mathbf{s} \tag{13}$$

The minimization of (13) subject to (12) can be solved by using Lagrange multipliers. It can be easily shown that the optimal coefficients of the unbiased p-channel marginal L-filter are given by:

$$\begin{aligned} \mathbf{a}^*_{(1)} &= \frac{s_1}{\underline{\hat{\mu}}_p^T \hat{\mathbf{R}}_p^{-1} \underline{\hat{\mu}}_p} \hat{\mathbf{R}}_p^{-1} \underline{\hat{\mu}}_p \\ \mathbf{a}^*_{(l)} &= \frac{s_l}{s_1} \mathbf{a}^*_{(1)} \quad l = 2, \ldots, p \end{aligned} \tag{14}$$

and the MMSE associated with the optimal coefficients (14) is:

$$(\varepsilon_{\text{unb}})_{\min} = \frac{1-\Delta}{\Delta} \mathbf{s}^T\mathbf{s} \tag{15}$$

If $\hat{\mathbf{R}}_p$ is positive definite, then $\Delta > 0$. It has been shown previously that $\Delta \leq 1$. Therefore, the MMSE associated with the optimal unbiased p-channel marginal L-filter is always greater than the MMSE (11) produced by the optimal unconstrained p-channel marginal L-filter. An estimate $\hat{\mathbf{s}}(k)$ of $\mathbf{s}$ has to be used in the design of the optimal unbiased multichannel marginal L-filter, too.

A multichannel marginal L-filter is said to be location-invariant, if its output is able to track small perturbations of its input, i.e., $\mathbf{x}'(k) = \mathbf{x}(k) + \mathbf{b}$ implies that

$$\mathbf{y}'(k) = \mathbf{T}[\mathbf{x}'(k)] = \mathbf{y}(k) + \mathbf{b} \tag{16}$$

where $\mathbf{y}(k) = \mathbf{T}[\mathbf{x}(k)]$. The definition of location-invariant multichannel marginal L-filter (16) yields the following set of constraints imposed on the filter coefficients:

$$\begin{aligned} \mathbf{e}^T\mathbf{a}_{jj} &= 1 \;\; \forall j, \;\; j = 1, \ldots, p \\ \mathbf{e}^T\mathbf{a}_{ji} &= 0 \;\; \forall i, \;\; i \neq j, \;\; j = 1, \ldots, p \end{aligned} \tag{17}$$

where $\mathbf{e}$ denotes the $(N \times 1)$ unitary vector, i.e., $\mathbf{e} = (1, 1, \ldots, 1)^T$. By incorporating (17) into (5) we obtain:

$$\varepsilon_{\text{loc}} = \sum_{i=1}^{p} \mathbf{a}^T_{(i)} \tilde{\mathbf{R}}_p \mathbf{a}_{(i)} \tag{18}$$

where

$$\tilde{\mathbf{R}}_p = \begin{bmatrix} \tilde{\mathbf{R}}_{11} & \tilde{\mathbf{R}}_{12} & \cdots & \tilde{\mathbf{R}}_{1p} \\ \tilde{\mathbf{R}}^T_{12} & \tilde{\mathbf{R}}_{22} & \cdots & \tilde{\mathbf{R}}_{2p} \\ \vdots & & \ddots & \vdots \\ \tilde{\mathbf{R}}^T_{1p} & \tilde{\mathbf{R}}^T_{2p} & \cdots & \tilde{\mathbf{R}}_{pp} \end{bmatrix} \tag{19}$$

with $\tilde{\mathbf{R}}_{ji} = \mathrm{E}[\tilde{\mathbf{n}}_j \tilde{\mathbf{n}}_i^T]$ $j, i = 1, \ldots, p$ and $\tilde{\mathbf{n}}_l = (n_{l(1)}, \ldots n_{l(N)})^T$, $l = 1, \ldots, p$. The minimization of (18) subject to (17) is formulated as minimization of the following Lagrangian function:

$$\Lambda(\mathbf{a}_{ji}, \lambda_{ji}; \; j, i = 1, \ldots, p) = \sum_{i=1}^{p} \mathbf{a}^T_{(i)} \tilde{\mathbf{R}}_p \mathbf{a}_{(i)}$$

$$+ \sum_{j=1}^{p} \{ \lambda_{jj}(1 - \mathbf{e}^T\mathbf{a}_{jj}) - \sum_{i=1, \; i \neq j}^{p} \lambda_{ji} \mathbf{e}^T \mathbf{a}_{ji} \} \tag{20}$$

Differentiating $\Lambda(\mathbf{a}_{ji}, \lambda_{ji};\ j,i = 1,\ldots,p)$ with respect to $\mathbf{a}_{ji}$ and equating the partial derivatives with zero, p independent sets of equations result, i.e.:

$$\breve{\mathbf{R}}_p \mathbf{a}^*_{(i)} = \frac{1}{2}\begin{bmatrix} \lambda_{1i}\mathbf{e} \\ \lambda_{2i}\mathbf{e} \\ \vdots \\ \lambda_{pi}\mathbf{e} \end{bmatrix} \quad i = 1,\ldots,p \tag{21}$$

which give the optimal coefficients $\mathbf{a}^*_{(i)}$ in terms of the Lagrange multipliers $\lambda_{1i},\ldots,\lambda_{pi}$. Let us assume that $\breve{\mathbf{R}}_p^{-1}$ exists and can be decomposed as follows:

$$\breve{\mathbf{R}}_p^{-1} = \begin{bmatrix} \mathbf{P}_{11} & \cdots & \mathbf{P}_{1p} \\ \mathbf{P}_{21} & \cdots & \mathbf{P}_{2p} \\ \vdots & \ddots & \vdots \\ \mathbf{P}_{p1} & \cdots & \mathbf{P}_{pp} \end{bmatrix} \tag{22}$$

where $\mathbf{P}_{ij},\ i,j = 1,\ldots,p$ are $(N \times N)$ square matrices. The Lagrange multipliers $\lambda_{1i},\ldots,\lambda_{pi}$ are obtained by solving the following set of equations:

$$\begin{bmatrix} \mathbf{e}^T\mathbf{P}_{11}\mathbf{e} & \cdots & \mathbf{e}^T\mathbf{P}_{1p}\mathbf{e} \\ \vdots & & \vdots \\ \mathbf{e}^T\mathbf{P}_{i1}\mathbf{e} & \cdots & \mathbf{e}^T\mathbf{P}_{ip}\mathbf{e} \\ \vdots & & \vdots \\ \mathbf{e}^T\mathbf{P}_{p1}\mathbf{e} & \cdots & \mathbf{e}^T\mathbf{P}_{pp}\mathbf{e} \end{bmatrix} \begin{bmatrix} \lambda_{1i} \\ \vdots \\ \lambda_{ii} \\ \vdots \\ \lambda_{pi} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 2 \\ \vdots \\ 0 \end{bmatrix} \tag{23}$$

or equivalently:

$$\lambda_{ji} = 2\frac{c_{ij}(\mathbf{G}_p)}{\det(\mathbf{G}_p)} \tag{24}$$

where $\mathbf{G}_p = \{G_{ij}\} \quad i,j = 1,\ldots,p$ is the left-hand side $(p\times p)$ square matrix of (23) and $c_{ij}(\mathbf{G}_p)$ stands for the cofactor of the ij-element of $\mathbf{G}_p$. In the following, the subscript p will be dropped out for notation simplicity. By using (21) and (24) the following optimal coefficients of the location-invariant p-channel marginal L-filter are obtained:

$$\mathbf{a}^*_{(i)} = \frac{1}{\det(\mathbf{G})}\breve{\mathbf{R}}_p^{-1}\begin{bmatrix} c_{i1}(\mathbf{G})\mathbf{e} \\ c_{i2}(\mathbf{G})\mathbf{e} \\ \vdots \\ c_{ip}(\mathbf{G})\mathbf{e} \end{bmatrix} \quad i = 1,\ldots,p \tag{25}$$

and the associated MMSE is given by:

$$(\varepsilon_{\text{loc}})_{\min} = \frac{1}{\det(\mathbf{G})}\sum_{i=1}^{p} c_{ii}(\mathbf{G}) \tag{26}$$

It is seen that the optimal coefficients (25) are independent of the two-channel constant signal to be estimated. Unfortunately, the location-invariant two-channel marginal L-filter leads only to a slightly higher noise suppression than its single-channel counterparts, as will be seen later on.

3 Computation of the moments of the multivariate order statistics

A discrete calculation of the correlation matrices and mean vectors is described in this section. It is based on the optimal quantization of each input signal (noise) vector component in the mean-squared-error sense. More specifically, discrete calculation of the correlation matrices and mean vectors is needed in order to avoid the extensive numerical integration involved in the definition of the moments of the order statistics [2, 8]. In the univariate case, a discrete calculation of the moments of the order statistics can be devised if each component of the input vector, which is a continuous random variable, is mapped into a discrete random variable. This can be accomplished by employing the optimum Lloyd-Max quantizer [6] whose design is well-established in the literature. It is also required to devise a discrete calculation of the moments of the order statistics in the bivariate case. To do so, the two-dimensional vectors of continuous random variables for all possible pairwise combinations between different channels should be mapped into two-dimensional vectors of discrete random variables by employing vector quantization. In our approach, numerical integration is required only in the design of Lloyd-Max quantizer, as will be seen later on.

Let us assume that $x_j,\ j = 1,\ldots,p$ has been quantized to M discrete values yielding a discrete random variable $x^*_j \in U_j = \{v_{j,1}, v_{j,2},\ldots,v_{j,M}\}$. Then, the elements of any submatrix of $\hat{\mathbf{R}}_p$, say $\mathbf{R}_{ji},\ j,i = 1,\ldots,p$ are given by:

$$\mathbf{R}^{rs}_{ji} = \mathrm{E}[x_{j(r)}x_{i(s)}] = \sum_{m=1}^{M}\sum_{n=1}^{M} v_{j,m}v_{i,n} \, f_{(r,s;j,i)}(v_{j,m}, v_{i,n}) \quad r,s = 1,\ldots,N \quad j,i = 1,\ldots,p \tag{27}$$

where $f_{(r,s;j,i)}(u_{j,m}, u_{i,n})$ can be evaluated in terms of $F_{(r,s;j,i)}(v_{j,m}, v_{i,n})$. For $i = j$, $F_{(r,s;j,j)}(v_{j,m}, v_{j,n})$ denotes the value at $(v_{j,m}, v_{j,n})$ of the cumulative distribution function (cdf) of the order statistics from a univariate population and well-known formulae [2, 8] can be applied. The above-mentioned formulae are in terms of $F_{x_j}(v_{j,m})$. Finally, the values of $F_{x_j}(v_{j,m})$ at $v_{j,m}$ are calculated by using the probabilities of discrete events $\{x^*_j = v_{j,q}\}$, i.e.:

$$F_{x_j}(v_{j,m}) = \sum_{q=1}^{m} \Pr\{x^*_j = v_{j,q}\} \quad m = 1,\ldots,M \tag{28}$$

The probabilities involved in (28) depend on the optimal decision levels $t_{j,q},\ t_{j,q+1}$ along the j-th dimension of the signal determined by the Lloyd-

Max quantizer design:

$$\Pr\{x_j^* = \upsilon_{j,q}\} = \int_{t_{j,q}}^{t_{j,q+1}} f_{x_j}(x_j)dx_j \qquad (29)$$

where $f_{x_j}(x_j)$ denotes the marginal pdf of the j-th random variable x_j.

For $i \neq j$, the results reported in [11, 13] can be exploited to evaluate the cdf values of the two-dimensional order statistic $(x_{j(s)}, x_{i(r)})$ at $(\upsilon_{j,m}, \upsilon_{i,n})$ in terms of the probability masses $\mathcal{F}_k(\upsilon_{j,m}, \upsilon_{i,n})$, $k = 0, \ldots, 3$ in the four regions of (x_j, x_i) plane. The probability masses can be easily calculated in terms of the probabilities of the discrete events $\{x_j^* = \upsilon_{j,m}, x_i^* = \upsilon_{i,n}\}$. The computation of these probabilities depends on the decision levels along the j-th and i-th dimensions provided by the Lloyd-Max quantizers and can be done without any difficulty.

The calculation of the elements of each mean vector $\underline{\mu}_j$, $j = 1, \ldots, p$ is performed by using:

$$\mathrm{E}[x_{j(r)}] = \sum_{m=1}^{M} \upsilon_{j,m} f_{(r;j)}(\upsilon_{j,m}) \quad r = 1, \ldots, N \qquad (30)$$

4 Multivariate Distributions

A multivariate distribution is said to be uniform, Gaussian, Laplacian etc. when the univariate marginal distributions are all uniform, Gaussian, Laplacian etc. [1, 4]. The previous attempts to use nonlinear filters based on order statistics for vector-valued signal processing either have been derived from a natural generalization of univariate exponential distributions [10, 9] or have been tested on a contaminated multinormal distribution which has been used to model long-tailed multivariate distributions [11]–[14]. In the following, the design of the bivariate Laplacian distribution is examined.

It is well-known [1, 4] that a joint distribution $F_{x_1,x_2}(x_1, x_2)$ given by:

$$F_{x_1,x_2}(x_1,x_2) = F_{x_1}(x_1)\, F_{x_2}(x_2)\, [1 + \alpha \times (1 - F_{x_1}(x_1))\,(1 - F_{x_2}(x_2))] \qquad (31)$$

where $\alpha \in [-1, +1]$, has as marginal cdf's $F_{x_i}(x_i)$ $i = 1, 2$. The family of joint distributions (31) is the so-called Morgenstern's family. We are interested in the case:

$$F_{x_i}(x_i) = \begin{cases} \frac{1}{2}\exp[\sqrt{2}\,\frac{x_i}{\sigma_i}] & \text{if } x_i < 0 \\ 1 - \frac{1}{2}\exp[-\sqrt{2}\,\frac{x_i}{\sigma_i}] & \text{if } x_i \geq 0 \end{cases} \qquad (32)$$

for $i = 1, 2$. This approach yields the bivariate Laplacian distribution.

Let $f_{x_1,x_2}(x_1, x_2; \theta_1, \theta_2) = f_{x_1,x_2}(x_1 - \theta_1, x_2 - \theta_2)$ where $\underline{\theta} = (\theta_1, \theta_2)^T$ is the location parameter of the distribution. For $\sigma_1 = \sigma_2 = \sqrt{2}$, the MLE of the location vector $\underline{\theta}$ based on a random sample of N observations $\{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$, where $\mathbf{x}_i = (x_{i1}, x_{i2})^T$ $i = 1, \ldots, N$, is the generalized vector median (GVM) [10] which uses the following distance function between the vectors $\mathbf{x}_j$ and $\mathbf{x}_i$:

$$\mathrm{dist}(\mathbf{x}_j, \mathbf{x}_i) = \|\mathbf{x}_j - \mathbf{x}_i\|_1 - \ln[1 + \alpha\, \mathrm{sgn}(x_{j1} - x_{i1}, x_{j2} - x_{i2})\,(1 - \exp[-|x_{j1} - x_{i1}|]) \times (1 - \exp[-|x_{j2} - x_{i2}|])] \qquad (33)$$

5 Simulation examples

A two-channel one-dimensional input signal sequence which obeys various bivariate distributions ranging from the short-tailed to the long-tailed ones has been used. In this experiment, $\mathbf{s}$ has been assumed known in order to compare quantitatively the performance of multichannel L-filters to that of their single-channel counterparts [8], as well as of vector median [10], of marginal median [11, 13] and of various ad-hoc estimators, such as the arithmetic mean. The quantitative criterion we used was the noise reduction index (NR) defined as the ratio of the output noise power to the input noise power, i.e.,:

$$\mathrm{NR} = 10 \log \frac{\sum_k (\mathbf{y}(k) - \mathbf{s})^T (\mathbf{y}(k) - \mathbf{s})}{\sum_k (\mathbf{x}(k) - \mathbf{s})^T (\mathbf{x}(k) - \mathbf{s})} \qquad (34)$$

Due to lack of space, only the performance of the nonlinear filters under study when a vector-valued constant signal $\mathbf{s} = (1.0, 2.0)^T$ is corrupted by additive white bivariate noise $\mathbf{n}(k)$ whose components are distributed according to the Laplacian-Morgenstern distribution is considered. A bivariate noise with zero-mean vector, $\sigma_1 = \sigma_2 = \sqrt{2}$ and $\alpha = 1.0$ has been assumed. The NR index is shown in Table 1 for filter lengths $N = 5$. It can be clearly seen that the unconstrained two-channel L-filter attains the highest noise reduction. The noise reduction capability of the unbiased two-channel L-filter approaches the one of the the unconstrained two-channel L-filter. Furthermore, the unbiased two-channel L-filter attains again almost 2 dB higher noise suppression than its single-channel counterpart [8]. This superior behavior is attributed to the fact that the unbiased two-channel marginal L-filter utilizes the correlation between the components of the input vector-valued signal. It is seen that the location-invariant two-channel marginal L-filter has only slightly better performance than its single channel counterpart. In general, the multichannel marginal L-filters have better performance than the arithmetic mean, the marginal median [11, 13] and the vector median [10]. The price which is paid for the superior performance is the complicated design procedure.

As can be seen from (10) and (14), the optimal coefficients of the unconstrained and unbiased p-channel marginal L-filters depend on the knowledge of the constant signal s (to be estimated) and of the joint pdf of the components of the input vector-valued signal $\mathbf{x}(k)$. In many practical applications, the constant signal s is unknown. The joint probability density function of the components of the input vector-valued signal $\mathbf{x}(k)$ can be easily estimated by its empiric joint pdf. Therefore, s must be estimated at every time-instant k from the past L-filter outputs $\mathbf{y}(l)$, $l = k-1, k-2, \ldots$. Let $\hat{\mathrm{s}}(k)$ denote the estimate of s at time-instant k. The estimate $\hat{\mathrm{s}}(k)$ will be used in the design of the optimal L-filters. These filters are applied to give the filter output $\mathbf{y}(k)$ which is a second (and hopefully better) estimate of $\mathrm{s}(k)$. The following estimation procedure for s is proposed:

$$\hat{\mathrm{s}}(k) = \frac{1}{N_e} \sum_{l=k-N_e}^{k-1} \mathbf{y}(l) \qquad (35)$$

where N_e is chosen to be sufficiently large. It has been verified by experiments that the estimator (35) having $N_e = 49$ is very good and gives the same noise reduction index as in the case of known s for a multichannel L-filter of length $N = 5$. The two-step procedure outlined above can be initialized by using the marginal median of the input vector-valued observations as an initial estimate of s.

Acknowledgments

The research reported in this paper has been supported by ESPRIT Basic Research Action project NAT (7130) and the project 89ED244 of the Greek Secretariat of Research and Technology.

References

[1] N.L. Johnson and S. Kotz, *Distributions in Statistics: Continuous Multivariate Distributions*. New York: Wiley, 1972.

[2] H.A. David, *Order Statistics*. New York: Wiley, 1981.

[3] A. Papoulis, *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill, 1984.

[4] L. Devroye, *Non-Uniform Random Variate Generation*. New York: Springer-Verlag, 1986.

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

[6] J. Max, "Quantizing for minimum distortion," *IRE Trans. Information Theory*, vol. IT-6, pp. 7-12, 1960.

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

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

[9] J. Astola and Y. Neuvo, "Optimal median type filters for exponential noise distributions," *Signal Processing*, vol. 17, pp. 95-104, 1989.

[10] J. Astola, P. Haavisto and Y. Neuvo, "Vector median filters," *IEEE Proceedings*, vol. 78, no. 4, pp. 678-689, April 1990.

[11] I. Pitas, "Marginal order statistics in color image filtering," *Optical Engineering*, vol. 29, no. 5, pp. 495-503, May 1990.

[12] L. Naaman and A.C. Bovik, "Least-squares order statistic filters for signal restoration," *IEEE Trans. on Circuits and Systems*, vol. CAS-38, no. 3, pp. 244-257, March 1991.

[13] 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. 247-260, September 1991.

[14] 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.

Table 1: Noise Reduction (in dB) for the Laplacian-Morgenstern Distribution Noise Model

Filter	NR ($N = 5$)
arithmetic mean	-6.966
marginal median	-7.642
vector median$_{L_1}$	-6.174
generalized vector median	-6.246
unconstrained 2-channel marginal L-filter	-12.774
unbiased 2-channel marginal L-filter	-12.617
location-invariant 2-channel marginal L-filter	-8.064
unbiased single-channel L-filter	-10.494
location-invariant single-channel L-filter	-8.046