Iterative LBG Clustering for SIMO Channel Identification

Fred Daneshgaran and Massimiliano Laddomada

Abstract: **This paper deals with the problem of channel identification for Single Input Multiple Output (SIMO) slow fading channels using clustering algorithms. Due to the intrinsic memory of the discrete–time model of the channel, over short observation periods, the received data vectors of the SIMO model are spread in clusters because of the AWGN noise. Each cluster is practically centered around the ideal channel output labels without noise and the noisy received vectors are distributed according to a multivariate Gaussian distribution. Starting from the Markov SIMO channel model, simultaneous maximum–likelihood estimation of the input vector and the channel coefficients reduce to one of obtaining the values of this pair that minimizes the sum of the Euclidean norms between the received and the estimated output vectors. Viterbi algorithm can be used for this purpose provided the trellis diagram of the Markov model can be labeled with the noiseless channel outputs. The problem of identification of the ideal channel outputs, which is the focus of this paper, is then equivalent to designing a Vector Quantizer (VQ) from a training set corresponding to the observed noisy channel outputs.**

The Linde-Buzo-Gray (LBG)-type clustering algorithms [1] could be used to obtain the noiseless channel output labels from the noisy received vectors. One problem with the use of such algorithms for blind time-varying channel identification is the codebook initialization. This paper looks at two critical issues with regards to the use of VQ for channel identification. The first has to deal with the applicability of this technique in general; we present theoretical results for the conditions under which the technique may be applicable. The second aims at overcoming the codebook initialization problem by proposing a novel approach which attempts to make the first phase of the channel estimation faster than the classical codebook initialization methods. Sample simulation results are provided confirming the effectiveness of the proposed initialization technique.

Index Terms: **Blind Equalization, clustering, channel estimation, fading.**

I. INTRODUCTION

Frequency-selective slow fading channels suffer from temporal dispersion which produces intersymbol interference (ISI) at the receiver. Diversity is an efficient means of counteracting the undesirable effects of the channel, which, in the presence of time-varying ISI, is modeled as a discrete-time finite memory system [2].

From a theoretical point of view, there are two main classes of algorithms which can be used for blind data detection. The first consists of evaluating an estimate of the inverse channel impulse response at the receiver and convolving this response with the received signal in order to detect the transmitted data. However, it is known that such algorithms suffer from noise enhancement, especially when channel introduces severe distortion characterized by deep nulls in the channel transfer function [2], [3]. The second approach accomplishes sequence detection via maximum likelihood estimation of the transmitted sequence. The recent literature have demonstrated that these algorithms could follow the channel dynamics in the presence of deep fading characterized by deep nulls in the channel impulse response [4]–[20]. Sato in [4] proposed a novel method to detect the transmitted data sequence without the knowledge of the channel response. In deep-fading situations, channel response must be identified rather than equalized. Sato presented algorithms that accomplish the tasks of channel identification and sequence detection without the use of a training sequence (i.e., in a blind mode). An excellent work related to the problem of sequence estimation for blind equalization has been conducted by Seshadri in [20]. In this paper, the author proposes a method for estimating the transmitted sequence using the maximum likelihood approach, by noting that the Viterbi algorithm (VA) could simplify the sequence estimation phase once the trellis diagram is labeled.

Wireless channels, in the presence of multipath fading, can be modeled using the hidden Markov model (HMM) [7]. This model is widely studied in the recent literature (see for example, [10]) and is so general that it could be adopted for describing other types of signal impairments, such as interference and non-Gaussian noise, in addition to fading and ISI whose effects extend over several symbol intervals.

The maximum likelihood sequence estimation is generally accomplished in two consecutive, or jointly interleaved steps, namely channel identification and sequence estimation. Kaleh et al. [6] proposed an iterative method for jointly conducting the channel estimation and the symbol detection via the Baum-Welch algorithm. Channel parameter estimation is accomplished through a maximum likelihood criterion. Blind Sequence estimation, without channel identification, has also been proposed by Tong [8], who suggested a method to estimate the source correlation from the observations of the received symbols without knowing the channel characteristics. Subsequently, VA is used to reconstruct the input symbols. In [9], a similar problem has been addressed. In comparison to the previous two works mentioned above, in [9], it is not assumed that the channel is static throughout the duration of the observed sequence and a time-varying channel response is adopted that the suggested method is able to track.

An approach based on maximum likelihood sequence estimation for signals over finite state Markov channels has been adopted by Kong et al. in [5]. In their work, the authors also proposed a maximum a posteriori probability(MAP) criterion to estimate the channel state sequence and the VA to obtain the optimum channel state sequence which is used in a second phase

Manuscript received February 1, 2003.

F. Daneshgaran is with the ECE Dept.,California State Univ., 5151 State University Drive, Los Angeles, CA 90032-8365, email: fdanesh@calstatela.edu.

M. Laddomada is with the Department of Electronics , Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy., email: laddomada@polito.it.

1229-2370/03/$10.00

to make soft-decision decoding. MAP criterion coupled with the expectation-maximization (EM) algorithm has been used by Turin in [10] whereby the EM algorithm is used to conduct MAP sequence estimation iteratively.

The problem of blind channel identification has been considered in [11], where a least-squares approach for SIMO channels has been proposed. In [12] and [13], the authors formulated a solution to the channel identification problem using second-order statistics following a time-domain and frequency domain approaches, respectively. In [14], the authors proposed a blind identification algorithm based on second-order statistics for single-input double-output channel model. Other algorithms proposed for blind channel identification have been presented in [15]-[16]. A review of blind identification algorithms proposed in the literature may be found in [17].

In connection with synchronization schemes for multipath channels, the authors propose, in [18], a nondata-aided carrier frequency offset estimator for noncircular (i.e., modulation formats with real-valued constellations) modulation formats for transmission through unknown frequency-selective channels. The advantage of the proposed algorithm derives from the fact that no knowledge of the multipath channel and transmitted data sequence is required, and that the synchronizer admits a feedforward structure that may be easily implemented using digital signal processing techniques.

In [19], the author considered the problem of blind symbol clock and frequency offset estimation in the context of time-selective fading channels, and compared the performance of various estimators based on the cyclic correlation properties of the signal. We shall assume throughout the rest of the paper that the carrier and symbol timing is acquired at the receiver and focus on the blind identification problem.

The assumptions stated above make the overall SIMO system resemble a vector Markov source whose noisy outputs constitute the observables. State transition analysis can be conducted on the Markov model with the aim of deducing the transition matrix of the model which could in turn be used to make simultaneous maximum-likelihood estimation of the input vectors and the channel coefficients. The problem then reduces to one of obtaining the values of this pair that minimizes the sum of the Euclidean norms between the received and the estimated output vectors using the Viterbi algorithm, provided the trellis diagram of the Markov model can be labeled with the noiseless channel outputs [21].

This paper focuses on the channel identification problem the solution of which may then be used in a more general sequence estimation problem, given the considerations stated above. In particular, this paper looks at two critical issues in the use of Vector Quantization (VQ) type clustering algorithms for channel identification. We aim to develop theoretical results that indicate conditions under which VQ techniques may be applicable for channel identification and develop an on-line initialization algorithm to speed up the codebook initialization process. Channel identification is then accomplished through LBG algorithm after initialization.

The paper is organized as follows. Section II discusses the application of clustering algorithms to channel identification, after developing the mathematical terminology adopted throughout the paper. Section III is devoted to the theoretical analysis of the applicability of VQ, followed by Section IV which presents the on-line initialization algorithm. Section V is devoted to simulation results verifying functionality of the proposed technique. Section VI aims at generalizing the obtained results to a more general class of channels suffering from ISI. Finally, in Section VII we present the conclusions.

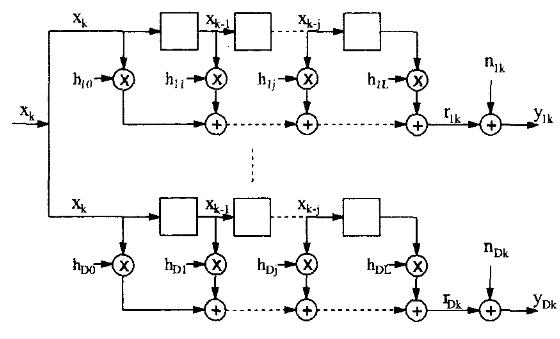


Fig. 1. SIMO channel model.

II. MATHEMATICAL FORMULATION OF THE PROBLEM

We start this section by establishing the terminology and notations used throughout this paper. Referring to Fig. 1, each branch in the SIMO model represents a distinct Finite Impulse Response (FIR) channel with memory L. We assume that every branch has the same memory dimension. The diversity order, that is the number of FIR branches in the SIMO model, is equal to D. The noise sequence corrupting the output of a given branch is an independent identically distributed (i.i.d.) zero mean AWGN sequence, and the noise sequences in different branches are independent of each other. We assume that the received signal in different channel branches undergo uncorrelated fading [2].

We assume that the transmitted symbols are binary PAM modulated, taking values from the set $\{-1, +1\}$ with equal probability. The extension to larger alphabet sizes and quadrature modulation is straightforward, but is omitted to keep the presentation lucid. The average received Signal to Noise Ratio (SNR) for branch d is:

$$\frac{\bar{E}}{N_o}|_d = \frac{E[|x_k|^2] \cdot \sum_{i=0}^{L} |h_{di}|^2}{N_o}, \tag{1}$$

and it is usually assumed that the SNR in each diversity branch is equal [3]. In the latter formula, $N_o = \frac{1}{2}E[|n_{dk}|^2]$, where n_{dk} is the k-th independent complex zero-mean white Gaussian noise samples added to the output of the d-th branch in the SIMO model. Since PAM modulation is assumed, actually only the in-phase component of the noise is of relevance, hence the factor N_o in the expression for SNR above. It is also assumed that $E[|x_k|^2] = 1$ [3].

With regards to the operation of the receiver, we assume that the carrier recovery loop generates a coherent carrier for performing the down conversion of the received RF signal. Samples of the matched filters' outputs taken at the symbol rate generate the matrix of sufficient statistic whereby each column is a vector whose components are the samples of the filter outputs at the receiver.

Our analysis is based on the polyphase decomposition of the output sequences of the SIMO channel depicted in Fig. 1. Such a decomposition allows for an easy description of the system dynamics as transitions between phases, and it is based on the following assumptions.

- The output of the SIMO channel is observed over an interval of length $B = K(L+1)$ symbols, with the starting time index set to 0, and the final time index set to $(K(L+1)-1)$. The input to the SIMO channel is the sequence $\vec{X} = \{x_j\}_{j=0}^{K(L+1)-1}$. Consider parsing the input sequence into groups of $(L+1)$ symbols. Let the l-th phase of the parsed input sequence be denoted by $\vec{X}^l = \{(x_{j(L+1)+l}, x_{j(L+1)+l+1}, .., x_{j(L+1)+l+L})\}_{j=0}^{K-1}$ for $l = 0, 1, ..., L$. Each $(L+1)$ tuple $(x_{j(L+1)+l}, x_{j(L+1)+l+1}, .., x_{j(L+1)+l+L})$ represents a super-symbol we denote by s_j^l. Hence, $\vec{X}^l = \{s_j^l\}_{j=0}^{K-1}$. The symbols $x_{K(L+1)}, x_{K(L+1)+1}, ..., x_{K(L+1)+L-1}$ are some fixed but arbitrary sequence unknown to the receiver.
- We partition the received symbols into a total of $(L+1)$ phases as well. In particular, the l-th phase of the i-th branch outputs is $\vec{Y}_i^l = \{y_{i,j}^l\}_{j=0}^{(K-1)} = \{y_{i,j(L+1)+l}\}_{j=0}^{(K-1)}$, $i = 1, 2, ..., D$ and $l = 0, 1, ..., L$, where $y_{i,t}$ denotes the t-th output of the i-th branch. In a totally analogous fashion, we may define the associated *noiseless* outputs $\vec{C}_i^l$ which denotes the l-th phase of the i-th channel noiseless outputs.
- The FIR impulse response of the i-th channel is the set of coefficients $\vec{H}_i = [h_{i,0}, h_{i,1}, ..., h_{i,L}]'$, $i = 1, 2, ..., D$, where $[.]'$ denotes the transpose operation.

We can suppress the dynamics of the SIMO system model depicted in Fig. 1 by looking at any given phase of the received vectors (i.e., the set of observables $\vec{Y}_i^l$ for $i = 1, 2, .., D$). Suppressing the model dynamics, the estimation of the noiseless channel output labels (i.e., the set of values that $\vec{C}_i^l$ can assume) is essentially a clustering problem with the Euclidian norm distance metric. Note that the noiseless channel output vectors $\vec{c}_k^l = \{c_{i,k}^l\}_{i=1}^D = \{\vec{H}_i' s_j^l\}_{i=1}^D$ can assume a finite set of values since s_j^l can assume at most $(q+1)^{(L+1)}$ possible values if input is $(q+1)$-ary. The clustering problem reduces to minimization of $\sum_{k=0}^{K-1} ||\vec{y}_k^l - \vec{c}_k^l||^2$. In particular, given the set of received vectors $\vec{y}_k^l$, the objective is to design a Vector Quantizer (VQ) via a proper selection of the reconstruction points with the objective of minimizing the overall codebook distortion. Note that the construction of the VQ codebook in this case does not require the knowledge of the SNR and in particular N_o.

A standard iterative technique for VQ codebook design is the LBG algorithm [1]. In particular, let $\{\vec{c}_1^p, \vec{c}_2^p, .., \vec{c}_{(q+1)^{L+1}}^p\}$ denote the estimated noiseless channel outputs at iteration p. To update the codebook, we first classify the received vectors $\vec{y}_k^l$ based on the minimum Euclidean norm. Let $C_p^l(i)$ denote the set of indices of the vectors $\vec{y}_k^l$ whose nearest neighbor is $\vec{c}_i^p$. Then we may write

$$V_p^l = \sum_i \sum_{k \in C_p^l(i)} ||\vec{y}_k^l - \vec{c}_i^p||^2, \tag{2}$$

where, V_p^l denotes the value of the cost function at iteration p. Now, consider minimization of V_p^l with respect to the reconstruction vector $\vec{c}_j^p$:

$$\nabla_{\vec{c}_j^p} V_p^l = \nabla_{\vec{c}_j^p} \sum_{k \in C_p^l(j)} ||\vec{y}_k^l - \vec{c}_j^p||^2 = \vec{0}, \tag{3}$$

$$\vec{c}_j^{p+1} = \sum_{k \in C_p^l(j)} \vec{y}_k^l, \tag{4}$$

where, $\vec{c}_j^{p+1}$ is the updated reconstruction vector at iteration $(p+1)$. Above is the well known centroid condition in the LBG algorithm. In practice, not considering the model dynamics, clustering can be conducted on *all* the received vectors and not those associated with any given phase, although theoretically speaking, the channel identification should utilize the system dynamics albeit via a much more complex estimation process.

III. RELATIONSHIP BETWEEN CLUSTER SIZE AND SNR AT THE SIMO RECEIVER INPUT

The received vectors at the channel output are distributed in clusters whose dimensions are related to the noise variances of the SIMO channel model. Since it has been assumed that the noiseless output vectors are corrupted by AWGN noise at the receiver, the received vectors are spread in clusters centered around the ideal channel outputs and are distributed according to a multivariate Gaussian distribution.

Knowledge of the SNR at the receiver provides a valuable side information about the size of the clusters. In particular, we can determine the effective radius of the hyper-sphere within which the received vectors may fall with a certain probability. The real part of the zero mean complex noise vector that adds to the noiseless channel output vector at time index k is $\overline{n}_k = [n_{1kr}, n_{2kr}, .., n_{Dkr}]'$. We focus on the real part of the complex noise quantities since the modulation assumed is PAM. For a given realization of the noise vector, the hyper-sphere surrounding each noiseless channel output vector has radius: $R = \sqrt{\sum_{d=1}^D n_{dkr}^2}$. As noted above, each complex noise component is zero mean and all the components are assumed to have equal variance $2N_o$, therefore the squared radius of the hyper-sphere, R^2 denoted Y, has a central Chi-square distribution with D degrees of freedom. The resulting Probability Density Function (PDF) is:

$$f_Y(y) = \frac{1}{N_o^D 2^{D/2} \Gamma(D/2)} y^{D/2-1} e^{-y/2N_o}, \tag{5}$$

where, $\Gamma(p)$ is the Gamma function. The Cumulative Distribution Function (CDF) has a closed form expression for even D, but in general, values of CDF for a normalized unit variance components can be found in tabulated form [22]. For a given fixed SNR per branch $\frac{E}{N_o}|_d$ with E normalized to unity, the noise variance N_o can be easily computed and used to obtain the effective squared radius of the hyper-sphere surrounding each noiseless channel output label, containing a certain fraction of the total probability (henceforth denoted as the $M\%$ containment probability squared radius). For instance, for $E = 1$ and

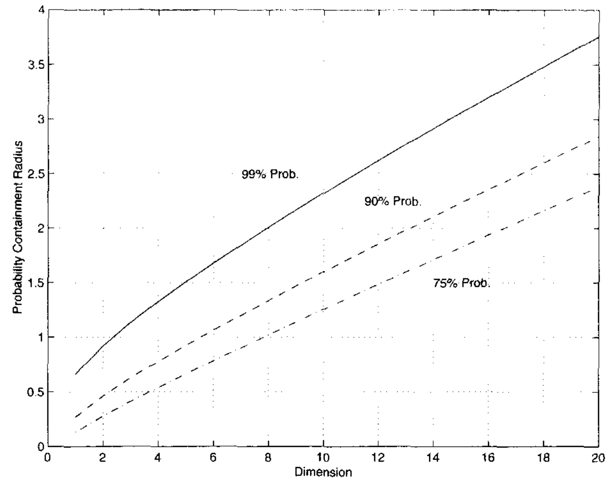


Fig. 2. Probability containment squared radii versus the number of diversity branches identifying the dimension of the observation space.

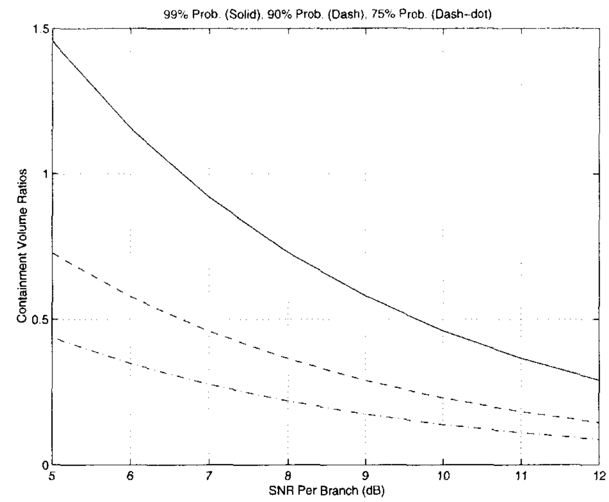


Fig. 3. Ratio of the volume of the noise to signal hyper-spheres for different containment probabilities as a function of the SNR per branch.

at an SNR of 10 dB per diversity branch, Fig. 2 depicts the 99%, 90% and 75% containment probability squared radii, for dimensions from 2 to 20. Clearly, as the number of diversity branches increase, the dimensionality of the observation space grows, and the probability containment squared radii grow as well. This growth in the containment squared radii should come as no surprise in light of the normalization that the SNR per branch of the SIMO model is constant. The implication being that as the number of branches increases, the cumulative received signal energy increases as well, a scenario that may or may not be true in practice. In what follows we shall consider two prevailing scenarios which may hold in practice, and for each draw conclusions about the applicability of clustering algorithms based on theoretical observations.

A. Spatial Diversity and Multiple Antenna Reception

In this scenario, several samples of the spatial domain signal are available using multiple antennas at the receiver. Hence, it is reasonable to assume that the signals at the output of different antennas have undergone uncorrelated fading and the mean SNR per branch of our SIMO model is constant.

From a clustering point of view, what matters is the fraction of the volume of the noise hyper-spheres for a given containment probability, relative to the total volume of the hyper-sphere containing all the noiseless channel output vectors (henceforth called the signal hyper-sphere). The volume of a hyper-sphere of radius r in n-dimension is given as:

$$Vol(r,n) = \frac{\pi^{n/2} r^n}{(n/2)!} \quad \text{for } n \text{ even}, \tag{6}$$

$$Vol(r,n) = \frac{2^n \pi^{\frac{n-1}{2}} (\frac{n-1}{2})!\, r^n}{n!} \quad \text{for } n \text{ odd}. \tag{7}$$

The squared radius of the signal hyper-sphere *per dimension* is one due to the normalization $E_{br} = 1$. Hence, the overall radius of the hyper-sphere in n dimension scales as $n^{n/2}$. Having the containment probability squared radii, obtained from the Chi-squared CDF tables, allows us to compute the ratio of the volume of the noise hyper-sphere for a certain containment probability, to the total signal hyper-sphere. This ratio provides a measure of the effectiveness of any clustering algorithm for channel identification. Fig. 3 depicts the volume ratio for 99%, 90% and 75% containment probabilities as a function of the SNR per branch of the SIMO system with two branches. As is evident from the figure, below 6.6 dB of SNR per branch, the 99% containment probability volume ratio is above one. While it is difficult to assess which containment probability should be used as a reference in assessing feasibility of clustering for channel identification, certain conclusions are consistent regardless of the actual value used. In particular, there is a threshold effect whereby below a certain SNR, the volume ratio exceeds one, implying that the noise hyper-sphere becomes greater than the signal hyper-sphere. In addition, note that under ideal circumstances, the number of noiseless channel output vectors is 2^{L+1} (assuming a binary input for the SIMO channel model). We note that this is an ideal circumstance because it is not necessarily

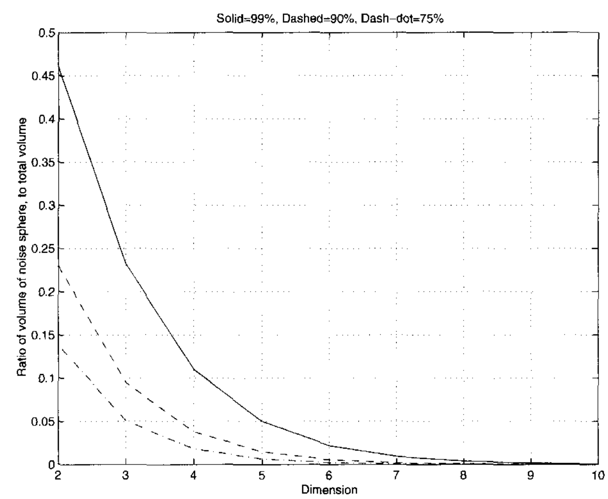


Fig. 4. Ratio of the volume of the noise to signal hyper-spheres for different containment probabilities at SNR of 10 dB per branch, as a function of the number of diversity branches.

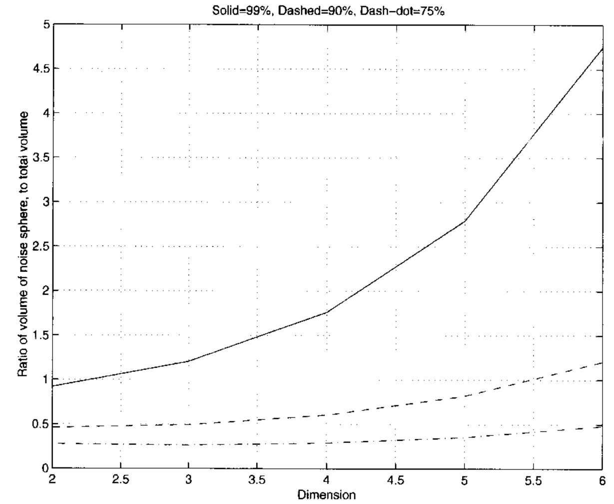


Fig. 5. Ratio of the volume of the noise to signal hyper-spheres as a function of the number of diversity branches for cumulative SNR of 10 dB at the receiver.

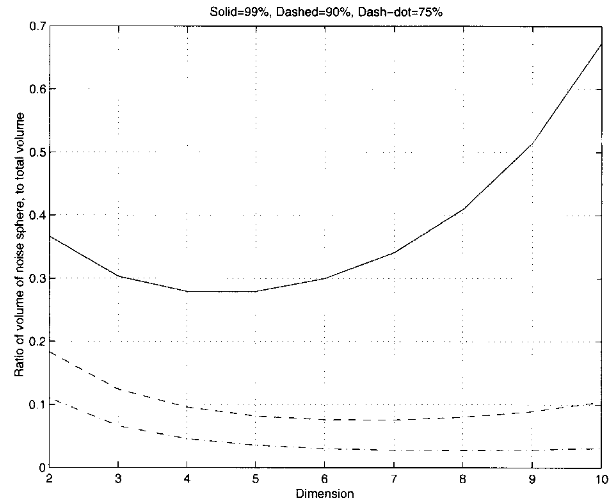


Fig. 6. Ratio of the volume of the noise to signal hyper-spheres as a function of the number of diversity branches for cumulative SNR of 14 dB at the receiver.

always true that all the possible noiseless output vectors are indeed distinguishable. There may be several binary combinations at the input leading to the same (or very close) noiseless output vectors. The conclusion being that the volume ratio above provides an upperbound on the number of noise clusters that can be packed in the signal hyper-sphere, and hence an upperbound on the allowed system memory L where clustering could be considered for channel identification. Based on the foregoing discussions, a general rule of thumb that may be applied regarding the SNR threshold is to consider the 75% containment probability volume ratio curve, and estimate the SNR at the point this ratio reaches about $2^{-(L+1)}$ (i.e., the packing density of the signal hyper-sphere under ideal circumstances). For a memory $L = 2$ system, the resulting SNR threshold is about 10 dB per branch.

The second issue concerns the choice of the number of diversity branches to use at the receiver. To assess this, Fig. 4 depicts the noise hyper-sphere to signal hyper-sphere volume ratios as a function of the number of diversity branches at a fixed SNR per branch equal to 10 dB. A key observation regarding this figure is that the volume ratio curves regardless of the containment probability, decay exponentially. The implication being that the knee of the exponential curve marks a distinct boundary point whereby most of the *gain* in using diversity for blind channel identification using clustering can be obtained. For our example at an SNR of 10 dB per branch, this corresponds to a SIMO with 5 branches. Note that this boundary point is consistent and independent of which containment probability value is actually used.

B. Frequency and/or Time Diversity Transmission and Reception

In this scenario, the total transmit power and the mean total received power is fixed regardless of the number of diversity branches (i.e., the total SNR is constant). Compared to the previous case, we can handle this scenario easily by scaling the SNR per branch based on the number of branches in the SIMO model.

In particular, let E_{tot} represent the total available signal energy at the receiver irrespective of the number of SIMO branches. Assuming equal SNR per branch, the resulting per branch SNR would be $E_{br}/N_o = (E_{tot}/D)/N_o = SNRb$. Now normalizing $E_{br} = 1$, we have that for a D-dimensional observation space with equal average SNR per diversity branch the noise variance per branch is $1/SNRb$. For instance, for a cumulative SNR of 10 dB, Fig. 5 depicts the ratio of the volume of the noise hyper-sphere to the signal hyper-sphere as a function of the number of diversity branches. Note that at 99% containment probability, the noise hyper-sphere almost occupies the entire volume of space for two diversity branches.

A threshold with respect to the use of a SIMO system for clustering can be observed as the cumulative SNR is varied. In particular, below about 10 dB cumulative SNR, the volume ratio curves grow with increasing number of diversity branches, indicating that below this value, use of a SIMO system with clustering applied for channel identification is not the best option. On the other hand above this threshold on the cumulative SNR, there is an optimal number of diversity branches where the volume ratio attains a minimum value. This optimal point is more or less the same regardless of the value of the containment probability itself (i.e., the minimum is broad for low values of containment probability). As an example, Fig. 6 depicts the ratio of the volume of the noise to signal hyper-spheres at a cumulative SNR of 14 dB for various containment probabilities. Note that the minimum ratio is achieved for $D \simeq 5$. We can estimate the optimal value of SNR per branch of the SIMO system as a function of the number of branches by looking at the curves of noise to signal volume ratios in D-space as a function of the number of SIMO branches and with SNR varied continuously. The result is reported in Table 1. Entries of Table 1 should not be interpreted as representing the only SNR values per branch of the SIMO system that could be used. Indeed as the SNR per branch of the SIMO system increases, the task of clustering is simplified as the effective noise radius for a given containment probability decreases. The numbers in essence provide the best trade-off between available SNR at the receiver and the number of SIMO branches in so far as the discernability of the noisy

D	SNRt (dB)	SNRb (dB)
2	11	8.0
3	12	7.2
4	13.5	7.5
5	14	7.0
6	15	7.2
7	15.5	7.0
8	16	7.0
9	16.2	6.7
10	16.5	6.5
11	16.7	6.3
12	17	6.2

Table 1. Optimal cumulative SNR and SNR per branch as a function of the number of SIMO branches, D is the number of SIMO branches, $SNRt$ is the total SNR and $SNRb$ is the SNR per branch.

clusters are concerned.

IV. CHANNEL IDENTIFICATION THROUGH VECTOR CLASSIFICATION AND CLUSTERING

In this section, starting from Equ. (3) and (4), we shall develop the channel identification procedure, simplifying the notations adopted in section II.

Use of the LBG algorithm [23] coupled with the knowledge of the sizes of the clusters allows data in each cluster to be used to obtain the respective cluster centroid, i.e. one of the ideal channel output vectors in absence of noise. Any clustering algorithm maps the k-dimensional received vectors into a finite set containing N centroids. In our problem, k is equal to the branch diversity at the receiver side, while the parameter N corresponds to the number of possible states derived from the channel model, i.e. $N = (q+1)^{L+1}$, where $q+1 = \#\{S\}$ is the cardinality of the source symbol ensemble and L is equal to the number of delay elements in the channel model. Clustering produces a codebook $C = \{\overline{c}_1, \overline{c}_2, \ldots, \overline{c}_N\}$ in which any vector $\overline{c}_i \in \Re^k$ for any $i \in J \equiv \{1, 2, \ldots, N\}$. Note that clustering attempts to associate each received vector $\overline{y}$ to the closest ideal channel output vector, that is the receiver output in absence of channel noise, in accordance to a suitable distance measure that will be specified later.

Generally speaking, whatever the adopted algorithm may be, the channel is uniquely identifiable only if it is possible to observe at the input of the receiver in the absence of noise, all the possible channel output labels. In other words, if the channel model has L delay elements and the transmitted symbols belong to an ensemble with cardinality $q+1$, it is necessary that all the $(q+1)^{L+1}$ possible labels are distinguishable at the input of the receiver. Otherwise, the channel cannot be uniquely identified.

The performance of the clustering algorithms are evaluated by means of a suitable distortion measure. In this paper, we have used a nonnegative cost function which determines the Euclidean distance between the ideal centroids and the ones reconstructed by the proposed algorithms.

Associated with any N point centroids is a partition of $\Re^k$ into N cells denoted R_i for $i \in J$. Each cell, or cluster, is identified by the following ensemble:

$$R_i \subset \{\overline{y} \in \Re^k : d(\overline{y}, \overline{c}_i) \leq d(\overline{y}, \overline{c}_j);\ \forall j \neq i\}, \qquad (8)$$

while the clustering algorithm applied to the problem of channel identification, could be represented as follows:

$$Q(\overline{y}) = \overline{c}_i \ only\ if\ d(\overline{y}, \overline{c}_i) \leq d(\overline{y}, \overline{c}_j)\ \forall j \neq i. \qquad (9)$$

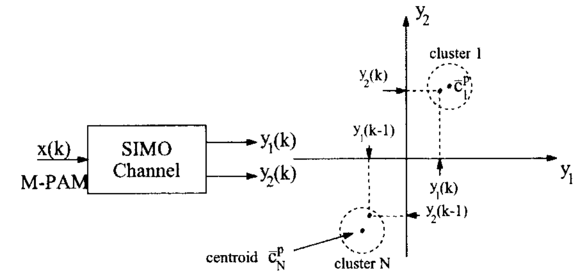


Fig. 7. Practical scheme of the vector clustering.

Clustering is fully determined by the codebook and the distortion measure as noted above. Note that, once the distortion measure is defined, no cell boundary needs to be specified in order to map the received vectors to the closest centroid in the codebook. If a received vector is equally distant from two or more codebook vectors, it is assigned to the first codebook vector already stored in the codebook.

The centroid condition is specified by the following relation:

$$\overline{c}_i = E(\overline{y} | \overline{y} \in R_i),\ where\ i \in \{1, \ldots, \#\{R_i\} = N\}. \qquad (10)$$

The formula above indicates that each centroid $\overline{c}_i$ corresponds to the mean of the vectors which fall in the cluster R_i (see Fig. 7). Conventional clustering algorithms, applied to source compression [23], use LBG-type algorithms to minimize, in successive iteration steps, the distortion between source symbols and the corresponding quantized symbols (i.e., the representative symbol) whose index is transmitted. Indeed, in this work we show that classical LBG algorithm applied to the received noisy channel outputs provides the noiseless channel output vectors. Details about the LBG algorithm applied to source compression could be found in [23].

A. Initialization and LBG Improvement of the Codebook

The core problem associated with the use of LBG type algorithms for time-varying channel identification is the codebook initialization which if not conducted properly, could lead to locally optimal codebooks. For our application this could essentially translate into having incorrect estimates of the noiseless channel outputs.

In this section, we deal with the problem of initializing the codebook vectors. Classical codebook initialization techniques such as Pruning, Random Coding, use of Product Codes and so on [23], create a codebook to which LBG iteration steps are then applied. However, these techniques requires one to scan all the data to be processed in order to generate N initial codebook vectors. Only after this phase, LBG algorithm is applied.

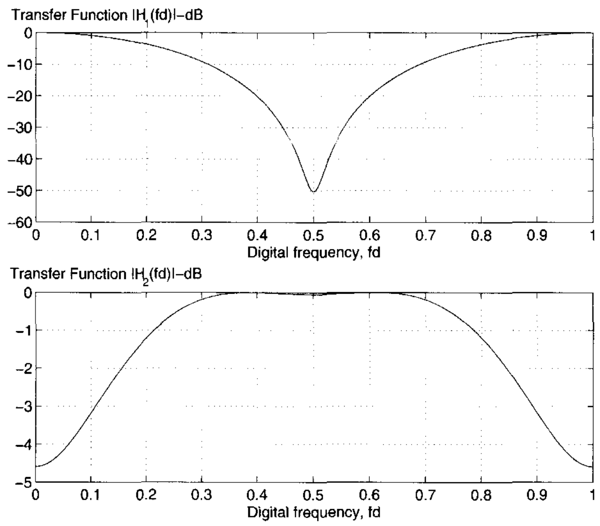


Fig. 8. Magnitude of the frequency responses of the 3-tap channels used in simulations.

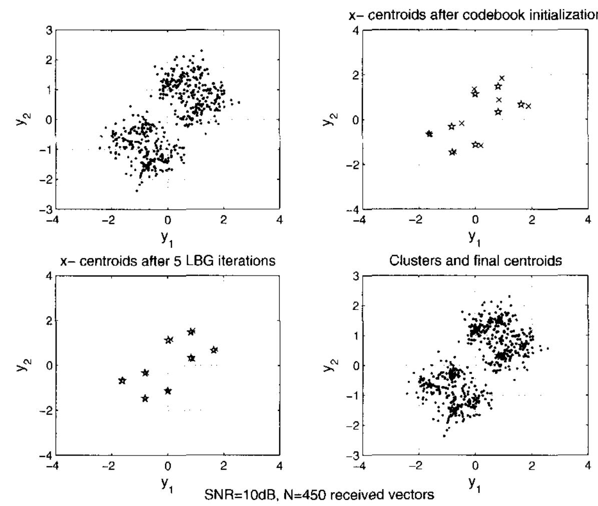


Fig. 9. SIMO channel clusters and related centroids obtained using the proposed initialization method.

In time-varying fading channels, it is very important that channel identification is conducted very quickly in order to allow the receiver to follow the channel dynamics. For this purpose, we propose a new method for creating the initial codebook. The technique allows one to create the codebook in an on-line fashion as soon as data are received. It works as follows; the first received data becomes a codebook vector. For each additional received data vector $\overline{y}_k = (y_{1k}, \ldots, y_{Dk})$, which we indicate as the vector received at the time index k (i.e. $\overline{y}_k$), the algorithm scans the vectors already inserted in the codebook and evaluates the Euclidean distances between them and the current data vector. If the minimum of these distances is greater than a predetermined threshold, the vector is inserted into the codebook as a new cluster centroid. Otherwise, the received vector is inserted in the cluster indicated by the closest vector already existing in the codebook. In the case the received vector is equally distant from two or more codebook vectors, it is associated to the first vector encountered. The mean value of all data belonging to a certain cluster will yield the cluster centroid according to Equation (10). This process is repeated for every received data vector. In this way, clusters and the codebook are evaluated in only one step. At the end of this first phase, the codebook vectors are replaced by those obtained via the centroid formula, that is:

$$\overline{c}_i^p = \frac{1}{n_i} \sum_{j=1}^{n_i} \overline{y}_{i,j} \ , \ \forall i = 1, .., N \ , \tag{11}$$

where p is the number of codebook improvement LBG iteration (in the first phase, $p = 1$), $\overline{y}_{i,j}$ indicates the j-th vector belonging to the i-th cluster, and n_i is the number of received vectors classified into the i-th cluster ($\sum_{i=1}^{N} n_i = B$, where B is the block size dimension, i.e., the number of received vectors on which clustering is performed). Note that received data are automatically classified into the various clusters based only on the distance properties.

In all simulations, the threshold thr has been chosen such that the clusters around the ideal centroids have a radius equal to $2\sigma_i$, $i = 1, \ldots, D$, where $\sigma_i = E[|n_i|^2]$. From Equ. (1), it is evident that all the σ_i, $i = 1, \ldots, D$ are equal.

Codebook improvement can be achieved via the application of the LBG algorithm. Starting from the initial codebook obtained using the procedure explained previously, we iteratively improve the codebook such that the codebook vectors converge toward the centroids of the channel output data clusters. After having classified the block of received data with respect to the centroids at step p, at the end of each LBG iteration new codebook centroids are evaluated using equation (11).

The codebook initialization technique proposed in this paper is compared to to the pruning method both in terms of speed of convergence, and the final distortion.

V. SIMULATION RESULTS

In this section, we present simulation results obtained for deep-fading SIMO channel with two branches having coefficients $\overline{h}_1 = [0.405, 0.817, 0.407]$ and $\overline{h}_2 = [0.17, 0.9, -0.4]$. The channel coefficients are for severely distorting channels [2] and satisfy the normalization $\sum_{i=0}^{L} |h_{ji}|^2 = 1$ for each branch $j = 1, 2$ (see Fig. 8). Note that the latter condition implies the use of automatic gain control in order for the variances of the channel outputs to be kept constant [24].

Simulation parameters are $D = 2$, $L = 2$, $2-$PAM binary signaling with source symbols $a_0 = -1$, $a_1 = +1$ and the SNR per branch is 10 dB. Simulation results for this SIMO channel model are shown in Fig. 9 and Fig. 10. The sub-figures located in the upper-left hand corner show all the received data vectors and the clusters they produce because of the effect of noise in the channel. In the upper-right hand corner of both figures, we depict the clusters obtained after the initialization stage with the proposed on-line algorithm (8 square points) and the initialization with the pruning technique. In the same figures, the ideal centroids are marked using stars. Note that the proposed algorithm results in centroids that are closer to the ideal noiseless channel output labels than those obtained using the pruning initialization technique. Figures located in the lower-left hand cor-

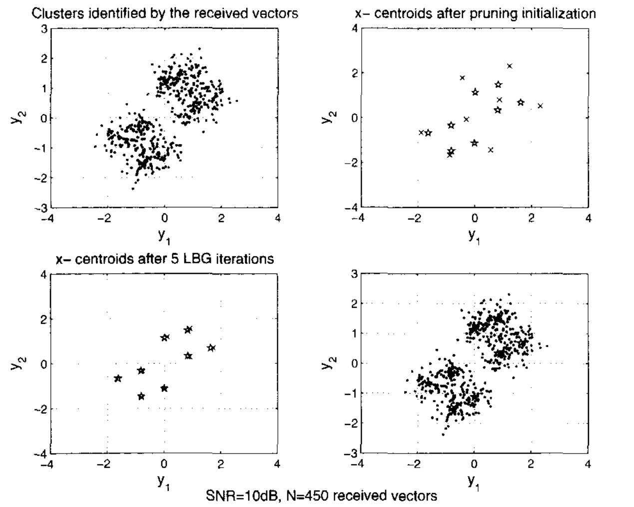


Fig. 10. SIMO channel clusters and related centroids in the case of pruning initialization.

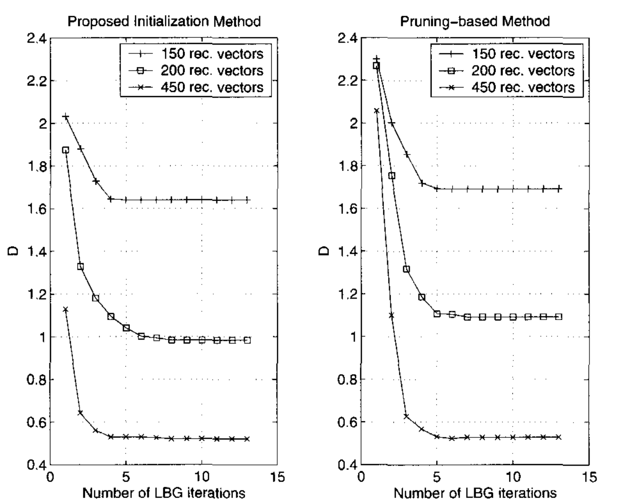


Fig. 11. Overall Distortion as a function of the number of LBG iterations.

ner show the centroids (square points) obtained after five LBG iterations, while the star points are the ideal centroids. In the right side of the same composite figure we show the clusters created by the received data and the centroids obtained after five LBG iterations. Note that the estimated centroids after five iterations are in the middle of the respective clusters. Other simulation parameters are also shown in the figures. Fig. 11 depicts the convergence behaviours of both techniques as a function of the number of LBG codebook improvement steps. Both figures have been parameterized with respect to the number of vectors on which clustering is performed, and show the overall distortion between the ideal and estimated centroids, evaluated using the following formula:

$$D = \sum_{j=1}^{N} d(\overline{c_j}^p, \overline{c_j}), \tag{12}$$

where, N is the number of clusters, $d(.,.)$ denotes the Euclidean distance between vectors, and $\overline{c_j}^p$ is the centroid $\overline{c_j}$ estimated at the p-th step of the LBG iteration.

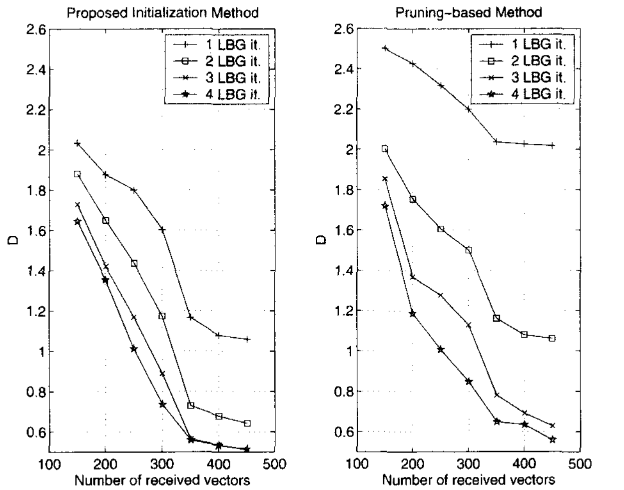


Fig. 12. Overall Distortion as a function of the number of the received vectors.

In Fig. 12, the same overall distortion has been plotted as a function of the LBG improvement step, and evaluated as a function of the number of the received vectors on which the LBG-based algorithms have been applied. Finally, Fig. 13 depicts the overall distortions as a function of the SNR at the receiver input.

Based on the simulation results, several observations can be made.

- Fig. 11 highlights that the convergence behaviour of the algorithms improves by increasing the number of the received vectors and the LBG steps after the initialization stage. However, with regards to the latter behaviour, after 5 LBG steps, overall distortions show asymptotic convergence that cannot be improved anymore. The same behaviours are observed in Fig. 12.
- Fig. 13 show that the overall distortion can be considerably improved by increasing the Signal-to-Noise ratio at the receiver input.
- All the figures show the effectiveness of using LBG-type algorithms to capture the dynamic variations of deep-fading channels and in particular the proposed initialization algorithm to speed up the convergence process of the estimated centroids toward the ideal ones.

VI. FURTHER RESULTS ON CHANNEL IDENTIFICATION

Generally speaking, channel identification through the proposed technique is effective in case the clusters are not overlapping. This problem, however, is related not only to the SNR at the receiver input, whose value affects the radii of the clusters and then the possibility for the clusters to overlap, but also on the number of ideal channel output labels, which depend on the dimension of the modulation scheme and the memory of the channel (i.e., the effects of ISI in producing adjacent symbols) impacting the relative distances between ideal centroids. The latter is strictly related to the channel characteristics. In the final analysis, SNR has the main influence. Even in the presence of deep nulls, the proposed methods are able to converge toward the ideal channel output labels using sufficiently high SNR val-

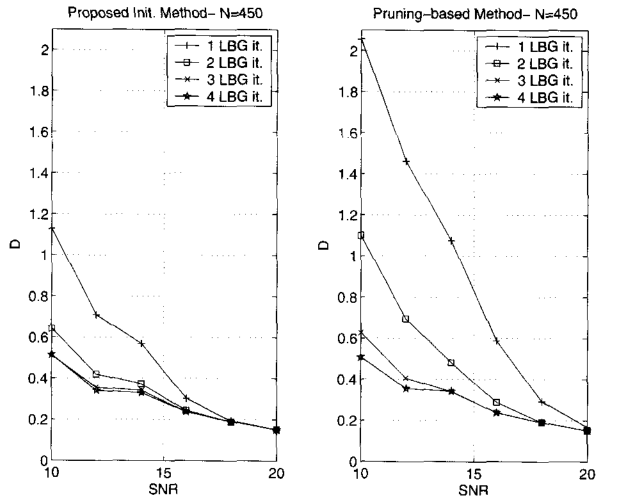


Fig. 13. Overall Distortion as a function of the Signal-to-Noise ratio at the receiver input.

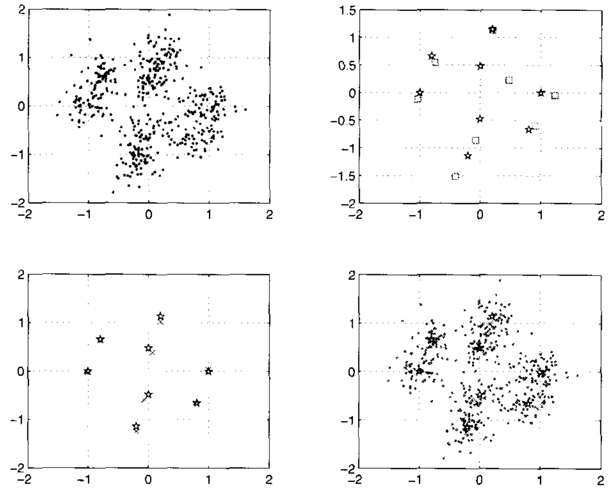


Fig. 14. Results obtained from the Simulation of the SIMO channel $< h1(n), h2(n) >$.

ues or via an increase in the number of diversity branches. The result is that the clusters formed by the received vectors would have effectively *more room* in the receiver space.

Many more simulations have been conducted on different channels whose results are not reported here for brevity, in order to verify the effectiveness of the proposed channel identification method. As an example, Fig. 14 shows the results obtained by applying the proposed initialization method and three steps of the LBG algorithm, to the SIMO $< h1(n), h2(n) >= (0.4,\ .5,\ .1),\ (-0.57,\ 0.24,\ 0.33)$. Note that the first channel, $h1$, has a frequency magnitude response with a deep null of -50 dB, while the second has a null of around -30 dB. Simulation parameters are as follows: $D = 2$, $L = 2$, $SNR = 10\ dB$. The number of received vectors on which clustering has been conducted is equal to 300 and the modulation format of the transmitted symbols is 2-PAM.

An effective mean of reducing the number of received vectors needed for accomplishing the clustering is to exploit the anti-symmetrical characteristic of the clusters identified at the receiver input. In other words, due to the symmetrical distribution of the M-PAM modulation, clusters are anti-symmetrically distributed in the space $(y1,\ y2)$. In this case, in fact, couple of ideal channel vectors differ in the sign, showing anti-symmetrical properties due to the type of signaling used. This deduction suggests one to perform clustering on a block of data containing only half of the number of data used in the method proposed before. Associating this approach to the on-line initialization described above, it is possible to obtain a fast channel identification process.

VII. CONCLUSIONS

This paper has focused on the problem of channel identification applied to slow-varying fading SIMO channels using clustering algorithms. The goal of the paper has been two folds; 1) to establish theoretical framework for determining when clustering may be used for channel identification and to develop insights about the role of the number of diversity branches and the received SNR for channel identification under different practical scenarios; and 2) to present an on-line LBG codebook initialization algorithm that leads to better convergence properties in comparison to other initialization techniques, both in terms of the rate of convergence and in terms of the overall distortion. Simulations have been used to show the effectiveness of the proposed algorithm.

REFERENCES

[1] Y. Linde, A. Buzo,and R. M. Gray, "An algorithm for vector quantizer design", *IEEE Trans. Commun.*, Vol. COM-28, No. 1, pp. 84-95, Jan. 1980.

[2] J. G. Proakis, "Digital communications", Mc Graw-Hill Publishers, 3-rd edition, 1995.

[3] G. L. Stuber, "Principles of mobile communication", Kluwer Academic Publisher, first edition, Jun. 1996.

[4] Y. Sato, "A blind sequence detection and its application to digital mobile communication", *IEEE J. on Select. Areas Commun.*, Vol. 13, No. 1, pp. 49-58, Jan. 1995.

[5] H. Kong, E. Shwedyk, "Sequence detection and channel state estimation over finite state markov channels", *IEEE Trans. Veh. Technol.*, Vol. 48, pp. 833-839, May 1999.

[6] G. K. Kaleh, R. Vallet, "Joint parameter estimation and symbol detection for linear or nonlinear unknown channels", *IEEE Trans. Commun.*, Vol. 42, pp. 2406-2413, Jul. 1994.

[7] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition",in *Proc.*, Vol. 77, No. 2, Feb. 1989, pp. 257-286.

[8] L. Tong, "Blind sequence estimation", *IEEE Trans. Commun.*, Vol. 43, No. 12, pp. 2986-2994, Dec. 1995.

[9] C. Anton-Haro et al., "On the inclusion of channel's time dependence in a hidden markov model for blind channel estimation", *IEEE Trans. Veh. Technol.*, Vol. 50, No. 3, pp. 867-873, May 2001.

[10] W. Turin, "MAP decoding in channels with memory, IEEE trans." Commun., Vol. 48, No. 5, pp. 757-763, May 2000.

[11] G. Xu et al., "A least-squares approach to blind channel identification", IEEE Trans. Signal Processing, Vol. 43, No. 12, pp. 2982-2993, Dec. 1995.

[12] L. Tong, G. Xu,and T. Kailath, "Blind identification and equalization based on second-order statistics: A time-domain approach", IEEE Trans. Inform. Theory, Vol. 40, No. 2, pp. 340-349, Mar. 1994.

[13] L. Tong et al., "Blind channel identification based on second-order statistics: A frequency-domain approach", IEEE Trans. Inform. Theory, Vol. 41, No. 1, pp. 329-334, Jan. 1995.

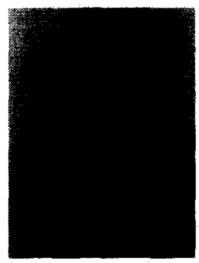
[14] Kimura et al., "Blind channel identification using RLS method based on second-order statistics," in *Proc.* IEEE Workshop on Signal Processing Advances in Wireless Communications, 1999, pp. 78-81.

[15] M. C. Vanderveen, A. Paulraj, "Improved blind channel identification using a parametric approach", IEEE Commun. Lett., Vol. 2, No. 8, pp. 226-228, Aug. 1998.

[16] Radionov, V., Mayrargue, S., "Semi-blind approach to second order identification of SIMO-FIR channel driven by finite-alphabet sequence", in *Proc.*, DSP 97., Vol. 1, Jul. 1997, pp. 115-118.
[17] L. Tong, S. Perreau, "Multichannel blind identification: From subspace to maximum likelihood methods", in *Proc.*, Vol. 86, No. 10, Oct. 1998, pp. 1951-1968.
[18] P. Ciblat, E. Serpedin,and Y. Wang, "On a blind fractionally sampling-based carrier frequency offset estimator for noncircular transmissions", IEEE Signal Processing Lett., Vol. 10, No. 4, pp. 89-92, Apr. 2003.
[19] Ghogho, M., Swami,and A., Durrani, T., "On blind carrier recovery in time-selective fading channels, conference record of the thirty-third asilomar conference on signals", Systems and Computers, Vol. 1, pp. 243-247, 1999.
[20] N. Seshardi, "Joint data and channel estimation using trellis search techniques", IEEE Trans. Commun., Vol. COM-42, No. 2/3/4, pp. 1000-1011, Feb./Mar./Apr. 1994.
[21] G. D. Forney, "Maximum likelihood sequence estimation of digital sequence in the presence of intersymbol interference", IEEE Trans. Inform. Theory, Vol. IT-18, pp. 363-378, May 1972.
[22] M. Abramowitz and I. A. Stegun, "Handbook of mathematical functions with formulas, graphs, and mathematical tables", 1964.
[23] A. Gersho, R. M. Gray, "Vector quantization and signal compression", Kluwer Academic Publisher, 1992.
[24] S. Haykin, "Adaptive filter theory", Prentice-Hall International, 2-nd edition, 1993.

Fred Daneshgaran received the B.S. degree in Electrical and Mechanical Engineering from California State University, Los Angeles (CSLA) in 1984, the M.S. degree in Electrical Engineering from CSLA in 1985, and the Ph.D. degree in Electrical Engineering from University of California, Los Angeles (UCLA), in 1992. From 1985 to 1987 he was instructor of the department of Electrical and Computer Engineering (ECE) at CSLA. From 1987 to 1993 he was assistant professor, from 1993 to 1996 he was an associate professor, and from 1997 he is a full professor of the ECE department at CSLA. From 1989 he has been the chairman of the communications group of the ECE department at CSLA.

From 1999 to 2001 he acted as the chief scientist for TechnoConcepts, Inc. where he directed the development of a prototype software defined radio system, managed the hardware and software teams and orchestrated the entire development process. In 2000, he co-founded EuroConcepts s.r.l. a R&D company specializing in the design of advanced communication links and software radio. Currently, he is acting as the chief executive officer and the chief technology officer of EuroConcepts s.r.l. In 1996 he founded Quantum Bit Communications, LLC., a consulting firm specializing in wireless communications. From 1992 he has been a research consultant to the TLC group of the EE department, Politecnico di Torino, Torino, Italy, where he consulted on a variety of projects for both national and European funded contracts, conducted joint research on wavelets, coded modulation, channel coding, etc. with members of the Signal Analysis and Simulation (SAS) group. Currently, he is serving as the editor of IEEE Transactions on Wireless Communications in the areas of modulation and coding, multirate and multicarrier Communications, broadband wireless communications, and software radio.



Massimiliano Laddomada was born in Modugno (Italy) in 1973. He received the degree in Electronics Engineering in 1999, and the Ph.D. degree in Communications Engineering from Politecnico di Torino, in 2003. He is currently Assistant Professor at Politecnico di Torino. From June 2000 to March 2001, he was a visiting researcher at CSLA (California State University), Los Angeles, and a consultant engineer in Technoconcepts, Inc., Los Angeles, a start up company specializing in Software Radio. He was awarded for his university career in Electronics Engineer with the fellowship by E.D.S.U. for five years on end. His research is mainly in wireless communications, especially modulation and coding, including turbo codes and, more recently, networks coding. Dr. Laddomada is a member of IEEE Communications Society and has co-authored over twenty papers in leading journals and international conferences. Currently, he is serving as a member of the editorial board of IEEE Communications Surveys and Tutorials. Dr. Laddomada is a contributor of TuttoScienze e Tecnologia, La Stampa, Italy.