



Multiple Composite Hypothesis Testing: A Competitive Approach

DAVID LUENGO, CARLOS PANTALEON, IGNACIO SANTAMARIA, LUIS VIELVA
AND JESUS IBÁÑEZ

Departamento de Ingeniería de Comunicaciones (DICOM), Universidad de Cantabria

Abstract. There are many problems in science and engineering where the signals of interest depend simultaneously on continuous and q -ary parameters, i.e. parameters which can take only one out of q possible values. This problem is generally known as multiple composite hypothesis testing. The probability function of the observed data for a given hypothesis is uncertain, as it depends on the parameters of the system. When there is no statistical model for the unknown continuous parameters, the GLRT is the usual criterion for the binary case. Although the GLRT philosophy can be extended to accommodate multiple composite hypotheses, unfortunately the solution is not satisfactory in the general case. In this paper, we restrict the general scenario and consider problems with q -ary input vectors and linear dependence on a unique set of continuous parameters; i.e. all the hypotheses depend on the same set of parameters. Direct application of the GLRT is feasible in this case, but it suffers from an exponential increase in complexity with data length. In this paper, we derive a low-complexity stochastic gradient procedure for this problem. The resulting algorithm, which resembles the LMS, updates the unknown parameters only along the direction of the winning hypothesis. This approach also presents similarities with competitive learning techniques, in the sense that at each iteration the different hypotheses compete to train the parameters. The validity of the proposed approach is shown by applying it to blind system identification/equalization, and chaotic AR(1) model estimation.

Keywords: multiple composite hypothesis testing, stochastic approximation, competitive training, blind system identification/equalization, chaos

Abbreviations: AR: Autoregressive; ARMA: Autoregressive Moving Average; AWGN: Additive White Gaussian Noise; CLMS: Competitive LMS; CRLB: Cramer-Rao Lower Bound; DOA: Direction of Arrival; FCLMS: Forced Competitive LMS; FIR: Finite Impulse Response; GLRT: Generalized Likelihood Ratio Test; HCLS: Hard Censoring Least Squares; i.i.d.: Independent Identically Distributed; ISI: Intersymbol Interference; LMS: Least Mean Squares; LS: Least Squares; LTI: Linear Time-Invariant; ML: Maximum Likelihood; PAM: Pulse Amplitude Modulation; PDF: Probability Density Function; PWL: Piecewise Linear; SNR: Signal to Noise Ratio

1. Introduction

Statistical inference problems in science and engineering can be grouped into one of two categories: detection and estimation. In the detection problem the objective is to select one out of a finite set of hypotheses according to some statistical criterion; i.e. select the possibility among a finite number of choices that is correct most of the times under some probabilistic measure. This problem is also known as hypothesis testing, and, when

there are more than two hypotheses, as classification or multiple hypothesis testing. In the estimation problem the objective is to determine the optimum value of a set of parameters of a signal; i.e. infer the values of the parameters trying to minimize some cost function based on a statistical model of the signal.

However, in many applications there are problems which require the simultaneous detection and estimation of several parameters. This problem is generally known as joint detection and estimation, or as multiple

composite hypothesis testing. Multiple composite hypothesis testing problems deal with the selection of one out of M hypotheses when the PDF under some or all the hypotheses is not completely specified. They appear in many different areas such as pattern recognition, time series analysis, or digital communications [1]. Typical problems in pattern recognition include optical character recognition [2], speech recognition [3], or machine vision and remote sensing [4]. Within the area of time series analysis, seismic deconvolution [5], detection and estimation of point processes [6], analysis of quasistationary and cyclostationary signals [7], or estimation of chaotic AR(1) signals [8] are problems that require both detection and estimation. Finally, problems in digital communications include estimation of signal parameters under uncertain signal presence [9], joint estimation of DOA and delay of a signal in a multipath environment [10], blind system identification/equalization [11], blind separation of multiple co-channel signals [12], or universal decoding in the presence of channel uncertainty [13].

Despite its importance and large family of applications, the problem of multiple composite hypothesis testing still lacks a satisfactory solution [14]. Nevertheless, a large variety of suboptimal algorithms have been developed, which can be divided in two categories [15]: statistical and non-statistical methods. On the one hand, statistical methods require a certain degree of knowledge of the statistical properties (PDF) of the observation process. On the other hand, non-statistical methods exploit particular features of the observation process [16]. Non-statistical methods are usually tuned to the specific class of signals for which they are devised, and are difficult to generalize.

In this paper we concentrate on statistical methods, where the unknown parameters can be modeled as random variables, or deterministic but unknown. In the first case the Bayesian approximation may be applied, integrating out the parameters to eliminate their influence in the detector [17]. This requires the a priori PDF of the parameters, which may not be known in many cases. In the second case the parameters are replaced by their ML estimates, leading to the GLRT [17]. Although the GLRT philosophy can be extended to accommodate multiple composite hypotheses, unfortunately the solution is not satisfactory in the general case; for instance, it is not able to discriminate hypotheses with nested parameter spaces [14]. However, in the problem considered, where all the hypotheses share the same set of parameters, this limitation of the GLRT does not apply.

In this paper, we consider problems with q -ary input vectors and linear dependence on a unique set of continuous parameters. Direct application of the GLRT is feasible in this case, but it suffers from an exponential increase in complexity with data length. To circumvent this problem, we derive a low-complexity stochastic gradient procedure. The resulting algorithm, which resembles the LMS, updates the unknown parameters only along the direction of the winning hypothesis. This approach also presents similarities with competitive learning techniques, in the sense that at each iteration the different hypotheses compete to train the parameters. The paper is organized as follows. In Section 2 the problem is described, and the mathematical notation is introduced. In Section 3, a stochastic gradient procedure for this problem is derived, relegating to the Appendix the cumbersome mathematical details. Finally, Section 4 shows the application of the proposed method in two scenarios: Blind Channel Identification/Equalization, and Chaotic AR(1) Model Estimation.

2. Problem Statement

The problem can be stated as follows: we have a sequence of samples of a signal of interest $x[n]$ observed in noise; i.e. our observations are

$$y[n] = x[n] + w[n], \quad n = 0, 1, \dots, N; \quad (1)$$

where $w[n]$ is a stationary, zero-mean AWGN process with variance σ^2 . The signal of interest is obtained as a linear transformation of a vector $\mathbf{v}[n]$, using a vector of unknown parameters $\boldsymbol{\theta}$:

$$x[n] = \boldsymbol{\theta}^T \mathbf{v}[n], \quad n = 0, 1, \dots, N. \quad (2)$$

The vectors $\mathbf{v}[n]$ are themselves generated from a q -ary vector $\mathbf{s}[n]$ (i.e. each of its components can only take one out of q possible values) according to a nonlinear mapping

$$\mathbf{v}[n] = T(\mathbf{s}[n]), \quad n = 0, 1, \dots, N; \quad (3)$$

where $\mathbf{s}[n]$ is an $L_s \times 1$ vector with components $s_i[n] \in \{l_1, \dots, l_q\}$.

Our objective is to detect the vectors $\mathbf{s}[n]$ and estimate the parameters $\boldsymbol{\theta}$ simultaneously. Since for each $x[n]$ there are $M = q^{L_s}$ possible $\mathbf{s}[n]$, and we have $N + 1$ samples, there are M^{N+1} different possible combinations of $\mathbf{s}[n]$ values. Hence, we can formulate

M^{N+1} possible hypotheses $\mathcal{H}_i, i = 0, \dots, M^{N+1} - 1$. Since we cannot specify a prior PDF for the unknown parameters of the system in many cases, we will model them as deterministic but unknown, and use the GLRT.

For two hypotheses, \mathcal{H}_0 and \mathcal{H}_1 , the GLRT decides \mathcal{H}_1 if [17]

$$\mathcal{L}_G = \frac{p(\mathbf{y}; \hat{\boldsymbol{\theta}}_1, \mathcal{H}_1)}{p(\mathbf{y}; \hat{\boldsymbol{\theta}}_0, \mathcal{H}_0)} > \gamma, \quad (4)$$

where \mathcal{L}_G is the likelihood ratio, $\mathbf{y} = \{y[0] \dots y[N]\}^T$ is the observation vector, γ is a certain threshold, and $\hat{\boldsymbol{\theta}}_i$ is the ML estimator of $\boldsymbol{\theta}$ for the hypothesis \mathcal{H}_i ; i.e. the one that maximizes $p(\mathbf{y}; \boldsymbol{\theta}, \mathcal{H}_i)$.

Unfortunately, the GLRT philosophy cannot be extended, in general, to the case of multiple composite hypothesis testing [14]. This is due to the fact that $p(\mathbf{y}; \hat{\boldsymbol{\theta}}_i, \mathcal{H}_i)$ increases as the complexity of the model (number of parameters) increases, thus biasing the detection rule towards the most complex hypotheses. Besides, when the parameter spaces of several hypotheses are nested, $p(\mathbf{y}; \hat{\boldsymbol{\theta}}_i, \mathcal{H}_i)$ is always maximum for the most general hypothesis, which is always chosen [14]. However, for the subset of problems that we are considering, there is a unique set of parameters $\boldsymbol{\theta}$. This implies that all the models share the same complexity, and hence the extended GLRT can be applied. In this case, the extended GLRT chooses the k -th hypothesis, where

$$k = \arg \max_i \{p(\mathbf{y}; \hat{\boldsymbol{\theta}}_i, \mathcal{H}_i)\}, \quad i = 0, \dots, M^{N+1} - 1. \quad (5)$$

According to our data model, (1), the PDF of the observation process is

$$p(\mathbf{y}; \boldsymbol{\theta}, \mathcal{H}_i) = \frac{1}{(2\pi\sigma^2)^{(N+1)/2}} \times \exp\left(-\frac{1}{2\sigma^2} \sum_{n=0}^N (y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n])^2\right); \quad (6)$$

where $\mathbf{v}_i[n]$ is the $\mathbf{v}[n]$ vector for the hypothesis \mathcal{H}_i , and sample time n . Hence, the ML estimate of $\boldsymbol{\theta}$ for the hypothesis \mathcal{H}_i is the value that maximizes (6), or, equivalently, the one which minimizes

$$J(\boldsymbol{\theta}; \mathcal{H}_i) = \sum_{n=0}^N (y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n])^2; \quad (7)$$

which can be solved using the pseudoinverse.

3. Competitive Estimator

The block estimator presented in the previous section requires obtaining M^{N+1} estimates of $\boldsymbol{\theta}$, and selecting the one with the minimum value of $J(\boldsymbol{\theta}; \mathcal{H}_i)$. This algorithm has an exponential computational cost, which prevents its application for large data records. Besides, for on-line problems the model has to be adapted in a sample by sample basis, which requires low-cost, fast estimation procedures. In these cases we propose a low-complexity stochastic gradient algorithm. The cost function that we wish to minimize is:

$$J_1(\boldsymbol{\theta}) = \sum_{n=0}^N (y[n] - \boldsymbol{\theta}^T \mathbf{v}_c[n])^2, \quad (8)$$

where the subindex

$$c = \arg \min_i \{e_i^2[n]\}, \quad i = 0, \dots, M - 1; \quad (9)$$

labels the best hypothesis for a given sample, and

$$e_i[n] = y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n], \quad i = 0, \dots, M - 1; \quad (10)$$

is the approximation error of each hypothesis. In this case, we can consider only the M possible hypotheses for each sample, instead of the M^{N+1} hypotheses of the GLRT. The best hypothesis \mathcal{H}_c for each sample is a discontinuous function of all the possible vectors $\{\mathbf{v}_0[n], \dots, \mathbf{v}_{M-1}[n]\}$, and the parameters $\boldsymbol{\theta}$. Therefore, a standard minimization of (8) is not possible. However, following a similar reasoning to [18], we can resort to the traditional procedure called stochastic approximation [19] to solve this problem. The basic idea is to use a sample function $e_c^2[n]$ of (8), and solve the problem using a local gradient-descent algorithm. Note that $e_c^2[n]$ is well defined and unique everywhere, except when the distance to two or more input vectors is exactly the same. Following this approach the algorithm becomes:

$$\begin{aligned} \boldsymbol{\theta}[n+1] &= \boldsymbol{\theta}[n] - \frac{1}{2} \mu[n] \nabla_{\boldsymbol{\theta}} e_c^2[n] \\ &= \boldsymbol{\theta}[n] + \mu[n] (y[n] - \boldsymbol{\theta}^T[n] \mathbf{v}_c[n]) \mathbf{v}_c[n]. \end{aligned} \quad (11)$$

The deterministic cost function in (8) can be considered a sample mean approximation of the more general

stochastic cost function

$$J_2(\boldsymbol{\theta}) = E\{(y[n] - \boldsymbol{\theta}^T \mathbf{v}_c[n])^2\}. \quad (12)$$

Since (12) is a discontinuous function of the vectors $\mathbf{v}[n]$, a standard minimization scheme is not possible, and we have to rely on iterative approaches. It may be shown (see the Appendix) that the exact gradient of (12) is

$$\nabla_{\boldsymbol{\theta}} J_2 = -2E\{(y[n] - \boldsymbol{\theta}^T \mathbf{v}_c[n])\mathbf{v}_c[n]\}. \quad (13)$$

This means that the steepest descent direction of (12) occurs in the direction of the best hypothesis (i.e. the winner), and leads to the CLMS algorithm:

$$\boldsymbol{\theta}[n+1] = \boldsymbol{\theta}[n] + \mu[n](y[n] - \boldsymbol{\theta}^T \mathbf{v}_c[n])\mathbf{v}_c[n]; \quad (14)$$

being \mathcal{H}_c the hypothesis that produces the minimum instantaneous output error $e_c[n]$. Note that (14) is identical to the stochastic gradient descent algorithm (11). The CLMS algorithm can be summarized as:

1. Initialize the parameter vector $\boldsymbol{\theta}$ with a random value.
2. Choose randomly an observation $y[n]$ and construct the M possible vectors $\mathbf{v}_0[n], \dots, \mathbf{v}_{M-1}[n]$ using the M possible q -ary vectors $\mathbf{s}_0[n], \dots, \mathbf{s}_{M-1}[n]$.
3. Test all the vectors $\mathbf{v}[n]$, select the one that provides the best match to the desired response $y[n]$ (i.e. least squared error) and use it to update the parameter vector $\boldsymbol{\theta}$ using (14).
4. Go to step 2 until stability is achieved.

This algorithm has been designed to solve a block estimation problem, but it can be easily adapted for on-line problems. In this case, with each new observation available, we construct a new set of M input vectors, which are applied to the system, using the best one to adapt the parameter vector $\boldsymbol{\theta}$.

The proposed procedure is related to the winner-takes-all competitive learning technique, which has been widely applied in the neural network literature [20, 21]. Clustering is probably the most important application of competitive learning. Given a collection of observations $\mathbf{y}[n]$, the objective of clustering is to find regions in the input space with high sample density, and then represent the input space by the centers of these regions (clusters). In clustering, we iteratively choose

one of the $\mathbf{y}[n]$ and compare it with all the reference vectors using some metric. The winner of this competition reduces its distance (in the reference metric) to the training vector. When stability is achieved, each reference vector represents a group (cluster) of the training data [20]. Competitive learning causes each of the reference vectors to concentrate on a particular group of patterns. This idea has been extended to models in [22], where several linear models compete for training patterns, concentrating each one in some group of them which share some kind of similarity. In our problem, the idea is slightly different: since we have a limited set of inputs we perform the competition in the input vectors using a unique model, and the winner (i.e. the input that produces the minimum error) is used to train the model.

4. Applications

4.1. Blind System Identification/Equalization

4.1.1. Problem Description and Previous Work.

Channel equalization is needed in bandlimited digital communications systems to compensate for the distortion caused by ISI. When a training sequence is not available (blind equalization/identification), and assuming without loss of generality an FIR model for the channel, this problem can be cast within the framework of multiple composite hypothesis testing. We consider a baud-rate sampled baseband representation of the digital communications system. A sequence of i.i.d. symbols $v[n]$ belonging to a finite alphabet is sent through an LTI channel with p coefficients $\theta[k]$. The resulting channel output is

$$y[n] = \sum_{k=0}^{p-1} \theta[k]v[n-k] + w[n], \quad (15)$$

where $w[n]$ is a zero-mean AWGN, and the signals and systems can be either real or complex. To simplify the discussion we concentrate on the real case, where we have a PAM input signal with $q = 2^Q$ levels $\{\pm 1, \pm 3, \dots, \pm(Q-1)\}$.

Assume we collect $N+1$ consecutive measurements, then $N-p+2$ equations according to (15) can be written in matrix form as

$$\mathbf{y} = \mathbf{V}\boldsymbol{\theta} + \mathbf{w}, \quad (16)$$

where the matrix \mathbf{V} contains the $N+1$ unknown input symbols. To find a joint solution for the best input

sequence and filter in (16), we should try all the M^{N+1} possible \mathbf{V} matrices (hypotheses), obtain for each one the minimum error norm solution for θ in (16) (which is the ML estimate for a given hypothesis), and select the filter and the corresponding matrix \mathbf{V} of input symbols which provide the best result. Obviously, the previous procedure corresponds to the extended GLRT stated in Section 2. However, due to its exponential computational cost, this approach is not practical. Note that, with this joint detection/estimation procedure, we obtain simultaneously an estimate for the channel and the input symbols; i.e. the channel is equalized. In this application, the nonlinear mapping $T(\cdot)$ becomes an identity, and thus $\mathbf{v}[n] = \mathbf{s}[n]$. On the other hand, the length of the q -ary discrete input sequence coincides with the length of the filter θ ; i.e. $L_s = p$.

In the context of blind equalization there are several solutions proposed to this problem. When the channel is known the best sequence can be obtained by applying the Viterbi algorithm [23]. When the channel is unknown several blind techniques involving also the Viterbi algorithm have been considered. In these techniques the channel is estimated in a decision-directed mode [24], or a set of "quantized" channel candidates are used to explore several trellises [25]. Other approach that does not rely on channel identification has been proposed by Tong in [26]: the source correlation is estimated from the observations and then a Viterbi algorithm is applied to estimate the input sequence. Several techniques that do not use the Viterbi algorithm have also been proposed. In particular, Yellin and Porat use an algebraic approach to identify an FIR system excited by a discrete-alphabet input [27]: they establish some conditions on the measurements from which the channel can be identified up to a sign ambiguity. Finally, in [28] only a fixed number of possible channels (those corresponding to the most likely sequence estimates) are evaluated.

4.1.2. Algorithms for Competitive Blind Equalization.

In this section the competitive approach previously described in a general framework is particularized for blind identification of FIR systems. A sample-by-sample strategy is considered, since it is more interesting in an equalization context. Two algorithms are proposed. First, for each new incoming sample, the M possible input sequences compete to model the new observation with the current filter estimate. The best input sequence is used to update the filter using the CLMS algorithm with a fixed adaption parameter. However,

the CLMS does not take into account that, at each time instant, only a new symbol has entered the filter and, therefore, the vectors of incoming symbols are correlated. This means that for each new incoming sample only q input sequences must compete for the new observation. These sequences are composed of the last $L_s - 1$ symbols of the previous winner and a new symbol belonging to the q -ary alphabet. This second algorithm is denoted as FCLMS. The algorithm is initialized by competing among the M possible hypotheses to fit the first observation, but, once a winner vector has been selected, only q hypotheses are considered for the rest of the observations. Therefore, the computational cost of the FCLMS is q^{L_s-1} times lower than that of the CLMS.

4.2. Simulation Results and Discussion

In this study we consider a binary ± 1 signal and two simple 2-tap channels: a minimum-phase channel $H_1(z) = 1 + 0.5z^{-1}$ and a nonminimum-phase channel $H_2(z) = 0.5 + z^{-1}$. In the first example we test the CLMS algorithm with channel $H_1(z)$. Figure 1 shows the evolution of the channel coefficients for 100 independent runs starting at random initial values in the interval $[-1.5, 1.5]$. The adaption parameter is fixed to $\mu = 0.002$, and the SNR is 10 dB. For this example, the mean value after convergence for both coefficients is one of the following four values ± 1 , or ± 0.5 . Therefore, we have eight possible filters $\pm 1 \pm 0.5z^{-1}$ and $\pm 0.5 \pm z^{-1}$, which are indistinguishable for the algorithm. The explanation for this identifiability problem is that, considering a noiseless situation, for each of the 8 possible filters and a given channel output, we can always find an input vector from the set $\{(-1, -1), (-1, +1), (+1, -1), (+1, +1)\}$ which produces that observation.

Applying the FCLMS algorithm to the same example, we obtain the results shown in Fig. 2. Now, forcing the correlation between consecutive winners, the filter converges to one of the two possibilities $\pm(1 + 0.5z^{-1})$. Hence, there is only a sign ambiguity, inherent to any blind equalization technique, which can be removed using differential coding. Repeating the simulation for the nonminimum-phase channel $H_2(z) = 0.5 + z^{-1}$, the CLMS algorithm provides similar results: depending on the initial condition, the algorithm converges to one of the eight possible filters $\pm 1 \pm 0.5z^{-1}$ or $\pm 0.5 \pm z^{-1}$, which are indistinguishable. On the other hand, for the FCLMS algorithm, in addition to the two global

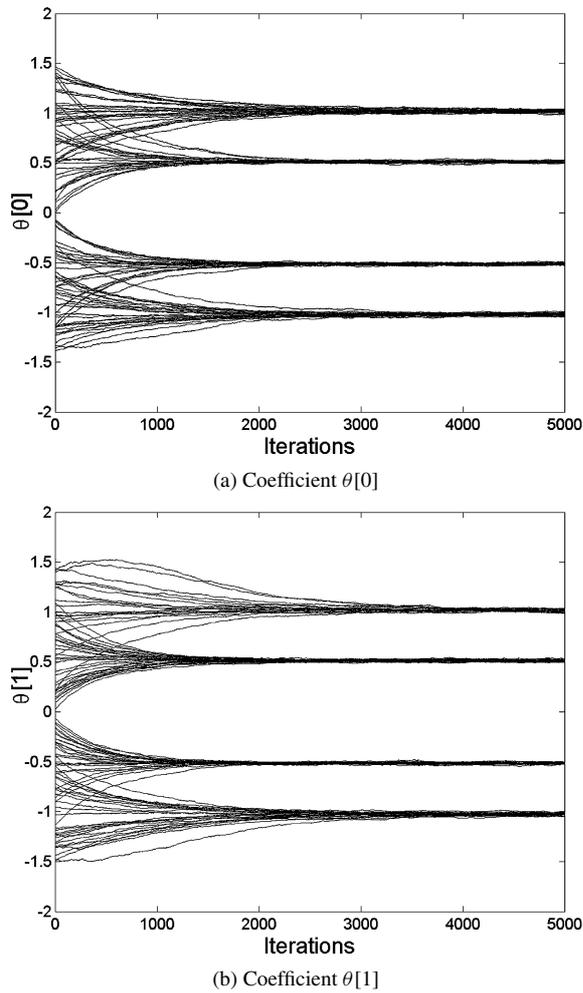


Figure 1. Convergence of 100 independent runs for the CLMS algorithm and channel $H_1(z) = 1 + 0.5z^{-1}$.

minima at $H(z) = \pm(0.5 + z^{-1})$, two strong local minima appear at $H(z) = \pm 1$, as is shown in Fig. 3.

In summary, for the CLMS algorithm and a filter of length L_s there are $L_s!2^{L_s}$ indistinguishable solutions, which are global minima of the cost function. In addition to these global minima some local minima can also appear. On the other hand, for the FCLMS algorithm we only have a sign ambiguity in the global minimum, but the population of local minima tends to increase. Therefore, the convergence of both algorithms to the desired global minimum strongly depends on an adequate filter initialization. This initial filter could be obtained, for instance, using the algebraic approach proposed in [27], or using a very short initial training sequence as occurs in packet data transmission or in semi-blind approaches [29]. Using this type of initial-

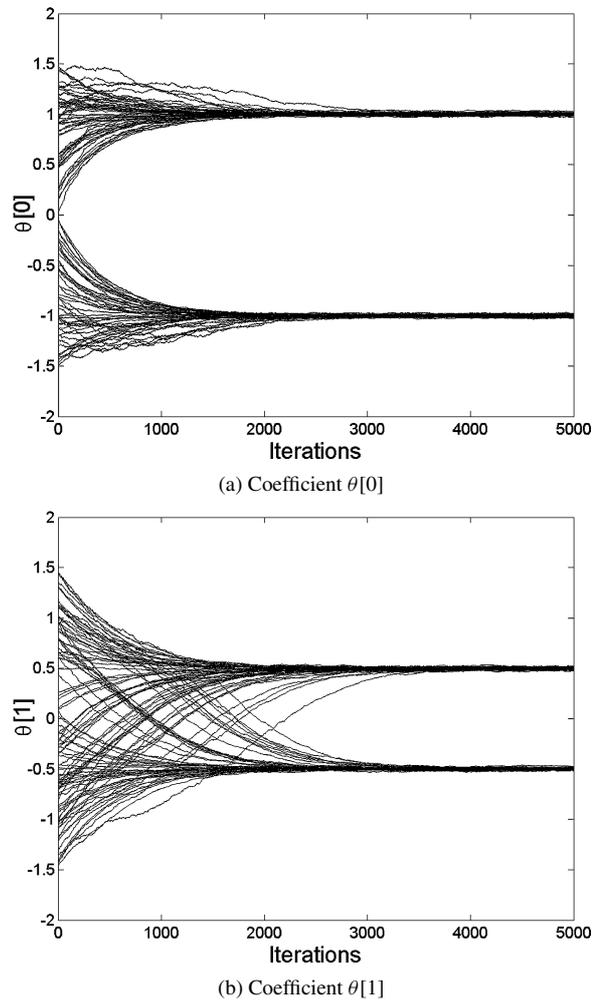


Figure 2. Convergence of 100 independent runs for the FCLMS algorithm and channel $H_1(z) = 1 + 0.5z^{-1}$.

ization, the proposed competitive approach is a very simple and effective approach to track a time-varying channel.

4.3. Estimation of AR(1) Chaotic Models

4.3.1. Problem Description and Previous Work.

Chaotic signals (i.e. signals generated by a nonlinear dynamical system in chaotic state) may be useful in modeling natural phenomena due to their special characteristics. For example, their extreme sensitivity to initial conditions makes signal generation a delicate task, but may be considered an advantage in representing anomalous behaviour of signals over short periods of time [30]. Chaotic models have been proposed for speech waveforms [30], biomedical signals [31], the

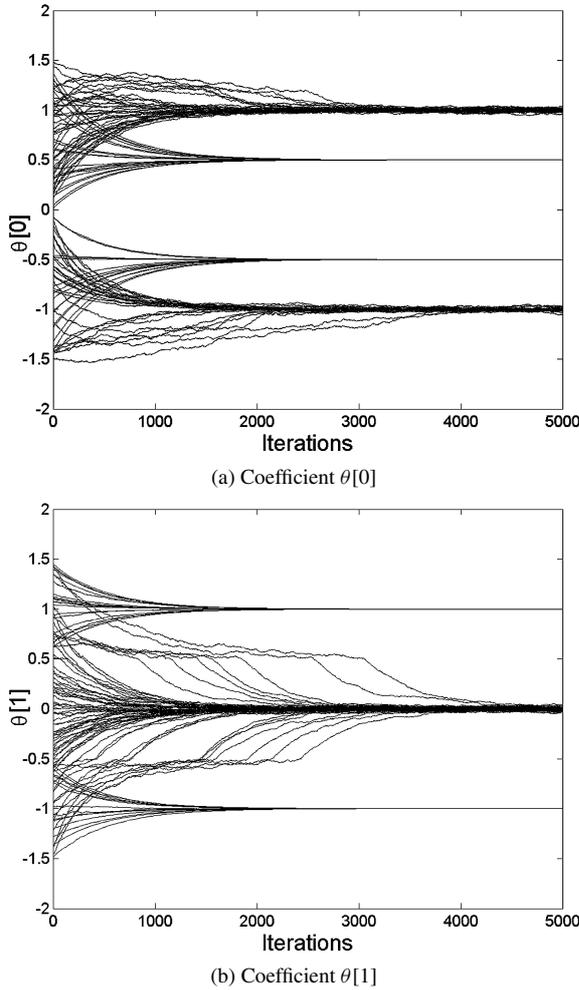


Figure 3. Convergence of 100 independent runs for the FCLMS algorithm and channel $H_2(z) = 0.5 + z^{-1}$.

sea clutter [32], packet traffic [33], as well as signals arising from many processes in experimental physics [34].

The application of chaotic modeling is conditioned by the lack of a family of chaotic models that combine a certain generality with easily computable estimation algorithms. Ideally, we would search for the chaotic equivalent of ARMA models. Chaotic signals generated by eventually expanding PWL Markov maps could be considered the chaotic equivalent of ARMA models, since they have rational spectra [35]. Nevertheless, it is unclear whether it is possible to construct chaotic PWL Markov maps with any desired spectra. Restricting the models to PWL maps also allows the analysis of ML signal estimators for a given map. The ML estimator is inconsistent, so the asymptotic distri-

bution for large data records is invalid. However, for a high SNR the ML estimator is asymptotically unbiased and attains the CRLB [36]. A closed-form expression for the ML estimator of chaotic signals generated by iterating known PWL maps has been derived in [37]. Parameter estimation has received much less attention, relying mostly on linear approaches, although ML estimators have also been considered in [38]. No closed-form solution is known for the ML estimator of the joint problem (parameter and signal estimator).

The chaotic signals that we are going to consider in this paper are generated according to

$$x[n] = F(x[n-1]), \quad (17)$$

where $F(\cdot)$ may be any nonlinear function. However, in this paper we concentrate on PWL maps, whose general expression is [39]

$$F(x) = \sum_{i=1}^K (a_i + b_i x) \chi_i(x), \quad (18)$$

where K is the number of disjoint convex intervals E_i in which the phase space of x may be divided, and $\chi_i(x)$ is an indicator function that denotes whether x belongs to the i -th interval or not

$$\chi_i(x) = \begin{cases} 1, & x \in E_i; \\ 0, & x \notin E_i. \end{cases} \quad (19)$$

In this paper we are going to use a particular family of PWL maps known as centered skew-tent maps, whose expression is

$$F(x) = \begin{cases} \frac{2(1+x)}{1+a} - 1, & x \leq a; \\ \frac{2(1-x)}{1-a} - 1, & x > a; \end{cases} \quad (20)$$

for some parameter $-1 < a < 1$. This family of maps can also be expressed using (18) with two intervals: $E_1 = [-1, a)$ with parameters $a_1 = 2/(1+a)$ and $b_1 = (1-a)/(1+a)$; and $E_2 = [a, 1]$ with parameters $a_2 = -2/(1-a)$ and $b_2 = (1+a)/(1-a)$.

This map produces sequences which are chaotic with invariant density $p(x)$ uniform in the range $[-1, 1]$ [40]. If a symbol from a known alphabet is assigned to each of the regions, the dynamics of the map may be characterized by following the different regions that the map visits during its dynamical evolution. This evolution is described by the sign sequence

$\mathbf{s} = s[0], \dots, s[N-1]$, also called itinerary, where

$$s[n] = i \Leftrightarrow x[n] \in E_i \quad (21)$$

Forward iteration of chaotic signals suffers from sensitive dependence on initial conditions. Therefore, we generate the map through backward iteration, i.e. $x[n] = F^{-1}(x[n+1])$. The centered skew-tent map is noninvertible, but, since it is unimodal, it has only two preimages, which can be found as

$$x[n] = 0.5[(1 - x[n+1])s[n] + a(1 + x[n+1])], \quad (22)$$

considering now a sign sequence given by

$$s[n] = \text{sign}(x[n] - a). \quad (23)$$

Note that it is necessary to know in advance the region to which the n -th sample belongs in order to generate it. The wide use of the family of skew-tent maps is due to the fact that their autocorrelation is [40]

$$R_{xx}[m] = r_0 a^m, \quad (24)$$

with $r_0 = 1/3$ in this case [8]. Therefore, the parameter a has the same relation with the autocorrelation as in the case of AR(1) processes, and skew-tent maps can be considered their chaotic equivalent.

4.3.2. Competitive Chaotic AR(1) Model Estimation.

In this section the general competitive framework is particularized for chaotic signals. In this case, the estimation of the model demands obtaining an estimate of the parameter a , the initial condition $x[0]$, and the itinerary \mathbf{s} . ML model estimation produces the initial condition and the parameter that minimize

$$J(x[0], a) = \sum_{n=0}^N (y[n] - F^{(n)}(x[0], a))^2. \quad (25)$$

This problem as it is stated has not been solved yet. However, the ML estimator is feasible, although of high computational cost. Minimizing (25) requires the computation of 2^N estimates, one for each possible itinerary, followed by the application of a gradient descent algorithm on a highly complex cost function. Therefore, an alternative algorithm based on the initial estimation of the parameter a using only pairs of samples, followed by an ML signal estimator has been considered [8].

To obtain an estimate of the parameter a we can exploit the deterministic relation that exists between

two consecutive samples, which leads to the following cost function using backward iteration:

$$J(a) = \sum_{n=1}^N (y[n-1] - F_{s[n-1]}^{(-1)}(y[n], a))^2. \quad (26)$$

Using (26) the problem may be decomposed into a set of linear ones as a function of the itinerary [8]. Obtaining the LS solution of (26) requires considering the 2^N possible itineraries and minimizing (26) for each one. However, in a moderate/high SNR situation (above 10 dB), it seems reasonable to consider only the $N+2$ possible itineraries produced by sorting the data samples and dividing them in two continuous sets. Thus, we obtain an HCLS estimate of the itinerary [8]. Once the parameter estimate has been obtained, we can apply the ML estimator in [37] to obtain the signal estimate.

Although the HCLS algorithm reduces the computational cost in the parameter stage from $\mathcal{O}(2^N)$ to $\mathcal{O}(N)$, for large data records or on-line estimation this algorithm is still impractical. In these cases, we can apply the CLMS algorithm described in Section 3. The parameter vector considered is $\boldsymbol{\theta} = [1 \ a]^T$, where the first parameter is fixed. The two possible input vectors $\mathbf{v}[n]$ at each iteration are obtained as

$$\begin{aligned} \mathbf{v}[n] &= \begin{cases} [-(1 - y[n])/2 & (1 + y[n])/2], & s[n] = -1; \\ [(1 - y[n])/2 & (1 + y[n])/2], & s[n] = 1; \end{cases} \end{aligned} \quad (27)$$

The competition and adaption is performed as described in Section 3 for both input vectors, with the particularity that in this case one of the parameters is fixed, and only the second coefficient of the filter needs to be adapted. Hence, the algorithm in this case becomes

$$a[n] = a[n-1] + \mu[n]e[n](1 + y[n])/2, \quad (28)$$

where $e[n]$ is the error after each iteration:

$$e[n] = \begin{cases} e_1[n], & e_1^2[n] \leq e_2^2[n]; \\ e_2[n], & e_1^2[n] > e_2^2[n]; \end{cases} \quad (29)$$

and $e_1[n]$ and $e_2[n]$ is the error for each of the two models:

$$e_1[n] = y[n-1] - \left(\frac{(1 + a[n-1])(1 + y[n])}{2} - 1 \right), \quad (30)$$

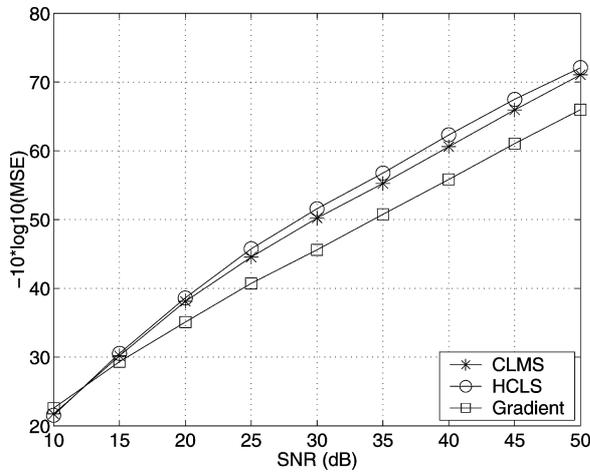


Figure 4. Comparison of alternatives of parameter estimation for $N = 99$, $a = 0.5$.

and

$$e_2[n] = y[n-1] - \left(1 - \frac{(1-a[n-1])(1+y[n])}{2}\right). \quad (31)$$

4.3.3. Simulation Results and Discussion. In this subsection we analyze the performance of the competitive chaotic model estimator, and compare it with the block model estimators developed in [8]. Concerning parameter estimation we compare the gradient descent approach and the HCLS solution with the CLMS algorithm. In Fig. 4 we show a typical MSE curve obtained averaging the results of 1000 simulations for each of the different values of SNR, for a skew-tent map with $a = 0.5$, $N = 99$ and a random initial condition. From Fig. 4 it can be inferred that the competitive method improves the performance of the gradient descent algorithm, and is very close to the HCLS estimate.

In this application our objective is to estimate as accurately as possible the chaotic signal. Therefore, to assess the performance of the proposed algorithm we evaluate the quadratic error of the whole sequence for $N = 99$ and different values of a and SNR. The parameter a is obtained using the CLMS algorithm, which provides a hard-censoring estimate of \mathbf{s} once convergence is attained. Then applying (22) recursively with the estimated itinerary and parameter, and using $x[N] = y[N]$, we may reconstruct the full sequence. The results are shown in Fig. 5. With this approach we achieve up to 20 dB of upgrade in SNR with respect

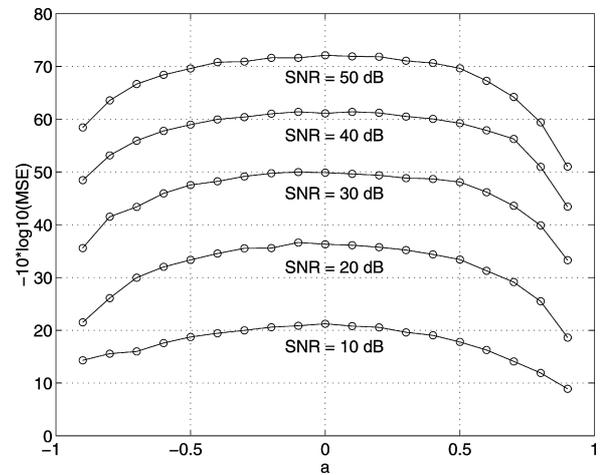


Figure 5. Mean Square Error of the estimated chaotic sequence for $N = 99$.

to the original signal, obtaining better results for low absolute values of the parameter a .

5. Conclusions

In this paper we have considered the multiple composite hypothesis testing problem, restricted to the case where all the hypotheses depend linearly on the same set of unknown parameters. Direct application of the GLRT to this problem is feasible, but shows an exponential increase in computational cost. As a low-complexity alternative, we have derived a stochastic gradient descent approach. Furthermore, we have shown that this algorithm can also be derived from a stochastic error functional (from which the cost function of the stochastic gradient is a sample function), where the exact gradient can be obtained. The resulting algorithm, which resembles the LMS, updates the unknown parameters only along the direction of the winning hypothesis. This approach also presents similarities with competitive learning techniques, in the sense that at each iteration the different hypotheses compete to train the parameters. The technique has been applied to the scenarios of blind system identification/equalization, and chaotic AR(1) model estimation.

Appendix

We follow a reasoning parallel to [18] to show the exact gradient descent algorithm for the optimization of the proposed cost function. The optimization of the cost

function $J_2(\boldsymbol{\theta}) = E\{e^2[n]\} = E\{(y[n] - \boldsymbol{\theta}^T \mathbf{v}_c[n])^2\}$ by gradient descent cannot be solved directly because $\mathbf{v}_c[n]$ is a discontinuous function of all the possible vectors $\{\mathbf{v}_0[n], \dots, \mathbf{v}_{M-1}[n]\}$, and the parameter vector $\boldsymbol{\theta}$. However, this problem can be circumvented by making use of the following lemma: if $\{a_i\}$ is a set of positive real scalar numbers, then

$$\min_i \{a_i\} = \lim_{r \rightarrow -\infty} \left(\sum_i a_i^r \right)^{1/r}. \quad (\text{A.1})$$

Another required result concerns the functional form

$$f(x, r) = (1 + x^{2r})^{1/r}. \quad (\text{A.2})$$

Excluding the values of x at which f or $\lim_{r \rightarrow -\infty} f$ are not differentiable, i.e. $x \in \{-1, 0, 1\}$, holds:

$$\lim_{r \rightarrow -\infty} \frac{\partial f}{\partial x} = \frac{\partial}{\partial x} \left(\lim_{r \rightarrow -\infty} f \right). \quad (\text{A.3})$$

To construct the gradient, notice that the function

$$\left[\sum_i (y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n])^{2r} \right]^{1/r}, \quad (\text{A.4})$$

is continuous, single-valued, well-defined, and continuously differentiable in its arguments, except when the error is zero or an input produces an error exactly equal to the sum of the rest. With a stochastic \mathbf{x} and continuous $p(\mathbf{x})$, all these singular cases have zero probability. Thus, under these conditions, and using the first result, the quadratic error $e[n]$ may be expressed as

$$\begin{aligned} e^2[n] &= (y[n] - \boldsymbol{\theta}^T \mathbf{v}_c[n])^2 \\ &= \min_i \{(y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n])^2\} \\ &= \lim_{r \rightarrow -\infty} \left[\sum_i (y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n])^{2r} \right]^{1/r}; \end{aligned} \quad (\text{A.5})$$

and, using the second result, the gradient becomes

$$\begin{aligned} \nabla_{\boldsymbol{\theta}}(J_2(\boldsymbol{\theta})) &= E \left\{ \lim_{r \rightarrow -\infty} \nabla_{\boldsymbol{\theta}} \left(\left[\sum_i (y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n])^{2r} \right]^{1/r} \right) \right\}. \end{aligned} \quad (\text{A.6})$$

Denoting

$$A = \sum_i (y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n])^{2r}, \quad (\text{A.7})$$

the gradient of the cost function becomes

$$\nabla_{\boldsymbol{\theta}}(J_2(\boldsymbol{\theta})) = E \left\{ \lim_{r \rightarrow -\infty} \nabla_{\boldsymbol{\theta}}(A^{1/r}) \right\}; \quad (\text{A.8})$$

where

$$\nabla_{\boldsymbol{\theta}}(A^{1/r}) = \frac{1}{r} A^{1/r-1} \nabla_{\boldsymbol{\theta}} A, \quad (\text{A.9})$$

and the gradient of A is

$$\begin{aligned} \nabla_{\boldsymbol{\theta}} &= r \sum_i [(y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n])^{2r-1}] \\ &\quad \times \nabla_{\boldsymbol{\theta}} [(y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n])^2] \\ &= -2r \sum_i [(y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n])^{2r-1}] \\ &\quad \times (y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n]) \mathbf{v}_i[n]. \end{aligned} \quad (\text{A.10})$$

Using (A.10), and after some straightforward rearrangements, we may express (A.9) as:

$$\begin{aligned} \nabla_{\boldsymbol{\theta}} A^{1/r} &= \\ &= -2A^{1/r} \frac{\sum_i [(y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n])^{2r-1} (y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n]) \mathbf{v}_i[n]]}{A}. \end{aligned} \quad (\text{A.11})$$

From (A.5)

$$\lim_{r \rightarrow -\infty} A^{1/r} = (y[n] - \boldsymbol{\theta}^T \mathbf{v}_c[n])^2. \quad (\text{A.12})$$

Denoting

$$\begin{aligned} B &= \frac{\sum_i [(y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n])^{2r-1} (y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n]) \mathbf{v}_i[n]]}{A} \\ &= \sum_i \frac{(y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n])^{2r} (y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n])^{-1} \mathbf{v}_i[n]}{\sum_j (y[n] - \boldsymbol{\theta}^T \mathbf{v}_j[n])^{2r}} \\ &= \sum_i \left(\sum_j \frac{(y[n] - \boldsymbol{\theta}^T \mathbf{v}_j[n])^{2r}}{(y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n])^{2r}} \right)^{-1} \\ &\quad \times (y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n])^{-1} \mathbf{v}_i[n], \end{aligned} \quad (\text{A.13})$$

and noticing that, when $r \rightarrow -\infty$, each of the terms in the inner sum for any given $\mathbf{v}_j[n]$, is maximum when

$\mathbf{v}_j[n] = \mathbf{v}_c[n]$, and starts to predominate progressively over the other terms:

$$\begin{aligned} \lim_{r \rightarrow -\infty} B &= \sum_i \lim_{r \rightarrow -\infty} \left[\frac{(y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n])^2}{(y[n] - \boldsymbol{\theta}^T \mathbf{v}_c[n])^2} \right]^r \\ &\quad \times (y[n] - \boldsymbol{\theta}^T \mathbf{v}_i[n])^{-1} \mathbf{v}_i[n] \\ &= (y[n] - \boldsymbol{\theta}^T \mathbf{v}_c[n])^{-1} \mathbf{v}_c[n]. \end{aligned} \quad (\text{A.14})$$

Combining the partial results, the exact gradient of the cost function becomes:

$$\nabla_{\boldsymbol{\theta}}(J_2(\boldsymbol{\theta})) = -2E\{(y[n] - \boldsymbol{\theta}^T \mathbf{v}_c[n])\mathbf{v}_c[n]\}. \quad (\text{A.15})$$

The sample function of the gradient at any instant n is

$$[\nabla_{\boldsymbol{\theta}}(J_2(\boldsymbol{\theta}))]_1 = -2(y[n] - \boldsymbol{\theta}^T \mathbf{v}_c[n])\mathbf{v}_c[n], \quad (\text{A.16})$$

and the steepest descent of J at instant n occurs in the direction of $-\nabla_{\boldsymbol{\theta}}(J_2(\boldsymbol{\theta}))_1$, leading to the CLMS algorithm:

$$\mathbf{v}_c[n+1] = \mathbf{v}_c[n] + \mu[n](y[n] - \boldsymbol{\theta}^T \mathbf{v}_c[n])\mathbf{v}_c[n]. \quad (\text{A.17})$$

Acknowledgments

This work has been partially financed by the European Commission and the Spanish Government under Grants 1FD97-1066-C02-01, and TIC 2001-0751-C04-03.

References

1. M. Feder and N. Merhav, "Universal Composite Hypothesis Testing: A Competitive Minimax Approach," *IEEE Trans. on Information Theory*, vol. 48, no. 6, 2002, pp. 1504–1517.
2. C.C. Tappert, Y. Suen, and T. Wakahara, "The State-of-the-Art in On-Line Hand-Written Recognition," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12, no. 8, 1990, pp. 787–808.
3. L.R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Proceedings IEEE*, vol. 77, no. 2, 1989, pp. 257–286.
4. A. Neri, "Optimal Detection and Estimation of Straight Patterns," *IEEE Trans. Image Processing*, vol. 5, no. 5, 1996, pp. 787–792.
5. J.M. Mendel, *Maximum-Likelihood Deconvolution: A Journey into Model-Based Signal Processing*, NY: Springer-Verlag, 1990.
6. C. Andrieu, E. Barat, and A. Doucet, "Bayesian Deconvolution of Noisy Filtered Point Processes," *IEEE Trans. on Signal Proc.*, vol. 49, no. 1, 2001, pp. 134–146.
7. M. Basseville and A. Benveniste, *Detection of Abrupt Changes in Signals and Dynamical Systems*, Berlin: Springer-Verlag, 1980.
8. C. Pantaleon, D. Luengo, and I. Santamaria, "Chaotic AR(1) Model Estimation," *Proc. ICASSP 2001*, Salt Lake City, UT (USA), vol. 6, 7–11 May 2001, pp. 3477–3480.
9. B. Baygun and A.O. Hero III, "Optimal Simultaneous Detection and Estimation Under a False Alarm Constraint," *IEEE Trans. on Info. Theory*, vol. 41, no. 3, 1995, pp. 688–703.
10. M. Wax and A. Leshem, "Joint Estimation of Time Delays and Direction of Arrival of Multiple Reflections of a Known Signal," *IEEE Trans. Signal Processing*, vol. 45, no. 10, 1997, pp. 2477–2484.
11. J.G. Proakis, *Digital Communications*, NY: McGraw-Hill, 1995.
12. A. Kannan and V.U. Reddy, "Maximum Likelihood Estimation of Constellation Vectors for Blind Separation of Co-Channel BPSK Signals and Its Performance Analysis," *IEEE Trans. on Signal Proc.*, vol. 45, no. 7, 1997, pp. 1736–1741.
13. J. Ziv, "Universal Decoding for Finite-state Channels," *IEEE Trans. Information Theory*, vol. IT-31, no. 4, 1985, pp. 453–460.
14. S.M. Kay, *Fundamentals of Statistical Signal Processing*, vol. II, *Detection Theory*, New Jersey: Prentice Hall, 1998.
15. G. Olmo, E. Magli, and L. Lo Presti, "Joint Statistical Signal Detection and Estimation. Part I: Theoretical Aspects of the Problem," *Signal Processing*, vol. 80, no. 1, 2000, pp. 57–73.
16. L.L. Scharf, *Statistical Signal Processing, Detection, Estimation and Time Series Analysis*, Reading, MA, USA: Addison-Wesley, 1991.
17. H.L. Van Trees, *Detection, Estimation and Modulation Theory*, New York, USA: Wiley, 1968.
18. T. Kohonen, "Self-Organizing Maps: Optimization Approaches," in *Artificial Neural Networks*, vol. II, Amsterdam: North Holland, 1991.
19. A. Hyvainen, J. Karhunen, and E. Oja, *Independent Component Analysis*, NY: John Wiley & Sons, 2001.
20. S. Haykin, *Neural Networks: A Comprehensive Foundation*, NY: MacMillan Pub. Co., 1994.
21. J. Principe, N.R. Euliano, and W.C. Lefebvre, *Neural and Adaptive Systems: Fundamentals Through Simulations*, NY: John Wiley & Sons, 2000.
22. C. Pantaleon, I. Santamaria, and A.R. Figueiras, "Competitive Local Linear Modeling," *Signal Processing*, vol. 49, no. 2, 1996, pp. 73–83.
23. G.D. Forney, "The Viterbi Algorithm," *Proc. of the IEEE*, vol. 61, no. 3, 1972, pp. 268–278.
24. F.R. Magee and J.G. Proakis, "Adaptive Maximum-Likelihood Sequence Estimation for Signaling in the Presence of Intersymbol Interference," *IEEE Trans. on Information Theory*, vol. IT-19, no. 1, 1973, pp. 120–124.
25. N. Seshadri, "Joint Data and Channel Estimation Using Blind Trellis Search Techniques," in *Blind Deconvolution*, Englewood Cliffs, NJ: Prentice Hall, 1994.

26. L. Tong, "Blind Sequence Estimation," *IEEE Trans. on Communications*, vol. 43, no. 12, 1995, pp. 2986–2994.
27. D. Yellin and B. Porat, "Blind Identification of FIR Systems Excited by Discrete-alphabet Inputs," *IEEE Trans. on Signal Processing*, vol. 41, no. 3, 1993, pp. 1331–1339.
28. F. Gustafsson and B. Wahlberg, "Blind Equalization by Direct Examination of the Input Sequences," *IEEE Trans. on Communications*, vol. 43, no. 7, 1995, pp. 2213–2222.
29. A. Gorokhov and P. Loubaton, "Semi-Blind Second Order Identification of Convolutional Channels," *Proc. ICASSP 1997*, Munich (Germany), vol. 5, 21–24 April 1997, pp. 3905–3908.
30. T.F. Quatieri and E.M. Hofstetter, "Short Time Signal Representation by Nonlinear Difference Equations," *Proc. ICASSP 1990*, Albuquerque, NM (USA), vol. 3, 3–6 April 1990, pp. 1551–1554.
31. M. Akay (Ed.), *Nonlinear Biomedical Signal Processing, Vol. II: Dynamic Analysis and Modeling*, New Jersey, USA: IEEE Press Series on Biomedical Eng., 2001.
32. S. Haykin and X.B. Li, "Detection of Signals in Chaos," *Proceedings of the IEEE*, vol. 83, no. 1, 1995, pp. 95–122.
33. P. Pruthi and A. Erramilli, "Heavy Tailed ON/OFF Source Behavior and Self-Similar Traffic," *Proc. ICC 1995*, Seattle, WA (USA), vol. 1, 18–22 June 1995, pp. 445–450.
34. T. Mullin (Ed.), *The Nature of Chaos*, NY: Oxford University Press, 1995.
35. S.H. Isabelle and G.W. Wornell, "Statistical Analysis and Spectral Estimation for One-Dimensional Chaotic Signals," *IEEE Trans. on Signal Processing*, vol. 45, no. 6, 1997, pp. 1495–1506.
36. S.M. Kay, "Asymptotic Maximum Likelihood Estimator Performance for Chaotic Signals in Noise," *IEEE Trans. on Signal Processing*, vol. 43, no. 4, 1995, pp. 1009–1012.
37. C. Pantaleon, D. Luengo, and I. Santamaria, "Optimal Estimation of Chaotic Signals Generated by Piecewise-Linear Maps," *IEEE Signal Processing Letters*, vol. 7, no. 8, 2000, pp. 235–237.
38. C. Pantaleon, D. Luengo, and I. Santamaria, "An Efficient Method for Chaotic Signal Parameter Estimation," *Proc. EUSIPCO 2000*, Tampere (Finland), vol. 3, 5–8 Sept. 2000, pp. 9–12.
39. D. Luengo, C. Pantaleon, and I. Santamaria, "Competitive Chaotic AR(1) Model Estimation," *Proc. IEEE XI NNSP Workshop*, North Falmouth, MA (USA), 10–12 Sept. 2001, pp. 83–92.
40. H. Sakai and H. Tokumaru, "Autocorrelation of a Certain Chaos," *IEEE Trans. on Signal Processing*, vol. 28, no. 5, 1980, pp. 588–590.



David Luengo was born in Santander, Spain, in 1974. He received the Radiocommunication Bachelor Engineer degree and the Telecomm Engineer degree from the Universidad de Cantabria, Spain, in 1994 and 1997, respectively. In 1997 he joined the Departamento de In-

geniería de Comunicaciones at the Universidad de Cantabria, Spain, where he is currently a Research Associate. His research interest include digital signal processing, digital communication systems and nonlinear systems.

david@gtas.dicom.unican.es



Carlos Pantaleón was born in Badajoz, Spain, in 1966. He received the Telecommunication Engineer degree and the Doctor degree from the Universidad Politécnica de Madrid (UPM), Spain, in 1990 and 1994, respectively. In 1990 he joined the Departamento de Ingeniería de Comunicaciones at the Universidad de Cantabria, Spain, where he is currently an Associate Professor. His research interests include digital signal processing, nonlinear systems and neural networks.

carlos@gtas.dicom.unican.es



Ignacio Santamaría was born in Vitoria, Spain, in 1967. He received the Telecommunication Engineer degree and the Doctor degree from the Universidad Politécnica de Madrid (UPM), Spain, in 1991 and 1995, respectively. In 1992 he joined the Departamento de Ingeniería de Comunicaciones at the Universidad de Cantabria, Spain, where he is currently an Associate Professor. His research interests include digital signal processing, nonlinear systems and neural networks.

nacho@gtas.dicom.unican.es



Luis Vielva was born in Santander, Spain, in 1966. He received the Licenciado in Physics and Doctor in Physics degrees from the Universidad de Cantabria, Spain, in 1997 and 1989, respectively. In 1990

he joined the Departamento de Ingeniería de Comunicaciones at the Universidad de Cantabria, Spain, where he is currently an Associate Professor. His research interest include digital signal processing, specially its applications to cosmology, biology and neuroscience.
luis@dicom.unican.es



Jesús Ibáñez was born in Santander, Spain, in 1971. He received the Radiocommunication Bachelor Engineer degree and the Telecomm

Engineer degree from the Universidad de Cantabria, Spain, in 1992 and 1995, respectively. In 1995 he joined the Departamento de Ingeniería de Comunicaciones at the Universidad de Cantabria, Spain, where he is currently an Associate Professor. His research interest include digital signal processing, digital communication systems and nonlinear systems.
jesus@gtas.dicom.unican.es