BLIND EQUALIZATION IN PRESENCE OF BOUNDED ERRORS

S. Icart, J. LeRoux, L. Pronzato, E. Thierry

Lab. I3S, CNRS-UNSA,
250 Av A. Einstein, 06560 Sophia Antipolis,
France

A.A. Zhigljavsky

Dept. of Math, St. Petersburg Univ.,
Bibliotechnaya sq.2, 198904,
Russia

ABSTRACT

This communication presents a new approach to blind equalization of a FIR channel. It is based on a bounded-error assumption and takes into account the fact that the input signal is in a finite alphabet. We show that even in the noisy case, identifiability can be guaranteed in finite time, provided that the support of the noise density is suitably bounded.

1. INTRODUCTION

The problem of equalization is of considerable interest in digital communication. Given a received output sequence, one has to determine the transmitted input sequence. If a training sequence is available, it permits to identify the channel, and then to recover the input data. To increase the capacity of the channel, it is of interest to remove this training sequence, which corresponds to blind equalization. This issue can be solved using High Order Statistics or Cyclic Statistics (see e.g. [2]). Instead of doing so, we choose to exploit the finite size of the input symbols set denoted by

$$\mathcal{L} = \{-(2M-1), \dots, -3, -1, 1, 3, \dots, 2M-1\}, \quad (1)$$

this real symmetric set being not restrictive, see [9].

Let us assume that the transmission channel is a Finite Impulse Response filter with known length p. The observation at time k is given by

$$y_k = \sum_{i=1}^{p} \bar{\theta}_i a_{k-i+1} + \epsilon_k = \boldsymbol{x}_k^T \bar{\theta} + \epsilon_k \quad (2)$$

with a_i the input symbols, $\bar{\theta} \in I\!R^p$ the unknown true value of the channel parameters, $\boldsymbol{x}_k$ the kth regressor,

$$\boldsymbol{x}_k = (a_k, a_{k-1}, \dots, a_{k-p+1})^T, \ a_i \in \mathcal{L} \ \ \forall i, \quad (3)$$

and ϵ_k an error term.

The equations (2) obtained for k between 1 and N can be put under the matrix form

$$\boldsymbol{y}_1^N = \boldsymbol{X}_1^N \bar{\theta} + \epsilon_1^N,$$

where $\boldsymbol{X}_1^N$ is a Toeplitz matrix containing the input sequence. Following the definition used in [1], we shall say that the input sequence is *identifiable* if it can be recovered up to a multiplicative scalar from the output sequence $\{y_k\}$.

In a noise free context, identifiability is shown in [1] to depend on the rank of $\boldsymbol{X}_1^N$. We show in this paper that identifiability can also be guaranteed *in finite time* in presence of noise, and with less restrictive assumptions, provided the support of the noise density is suitably bounded.

Paper is organized as follows : section 2 is devoted to estimation in presence of bounded error, section 3 to blind equalization in this case. Section 4 deals with identifiability. We then present some simulation before concluding.

2. BOUNDED-ERROR ESTIMATION

In data communication, the boundness of errors is generally not taken into account, although it is a realistic assumption. Let us assume that the errors ϵ_k are such that

$$\forall k, \ \|\epsilon_k\| \le \epsilon \le e, \quad (4)$$

where ϵ defines the (unknown) true support of the noise density, whereas e is a known upper bound for ϵ. The *likelihood set* associated with observations $\boldsymbol{y}_1^N$ and regressors $\boldsymbol{X}_1^N$ is then defined as

$$\mathcal{S}(\boldsymbol{X}_1^N) = \{\theta \in I\!R^p \ / \ |y_k - \boldsymbol{x}_k^T \theta| \le e, \ k = 1, \dots, N\}, \quad (5)$$

where the dependence of $\mathcal{S}$ in $\boldsymbol{y}_1^N$ is omitted.

Parameter bounding, which aims at characterizing the set $\mathcal{S}(\boldsymbol{X}_1^N)$, has received a growing attention in the last decade (see, e.g. [7, 5]). The situation considered here is not classical for parameter bounding in several aspects. First, since the alphabet $\mathcal{L}$ is finite, the regressors belong to a finite set,

$$\Phi = \{\phi_1, \dots, \phi_m\}\,,\ \phi_i \neq \phi_j \text{ for } i \neq j\,. \qquad (6)$$

This will greatly facilitate the characterization of the likelihood set. Second, the error bound e is not assumed to correspond to the exact support of the density of the ϵ_k's, and the convergence of $\mathcal{S}(\boldsymbol{X}_1^N)$ will be considered in the case $e \geq \epsilon$.

2.1. Construction of the likelihood set

Since the model structure (2) is linear with respect to the parameters, the likelihood set $\mathcal{S}(\boldsymbol{X}_1^N)$ given by (5) is a polytope (provided $\text{rank}(\boldsymbol{X}_1^N) = p$) that can be constructed *recursively*, see [8, 4]. At each step k, the vertices $\boldsymbol{v}$ that do not satisfy $y_k - e \leq \boldsymbol{x}_k^T \boldsymbol{v} \leq y_k + e$ are removed, and new ones are constructed along the edges joining a removed vertex with an adjacent one which is kept. Since the alphabet is finite, the complexity of $\mathcal{S}(\boldsymbol{X}_1^N)$ can be bounded. For instance, when $\mathcal{L} = \{-1, 1\}$ (that is when $M = 1$), $\mathcal{S}(\boldsymbol{X}_1^N)$ is a pallelotope with only $2p$ vertices, whatever the value of N (provided $\text{rank}(\boldsymbol{X}_1^N) = p$).

2.2. Convergence of the likelihood set

Contraction of $\mathcal{S}(\boldsymbol{X}_1^N)$ to a point (similar to the notion of consistency in classical point-estimation) is established in [6] under rather general hypotheses, assuming that $e = \epsilon$ and that ϵ reaches its bounds infinitely often (see H1 below). However, assuming $e = \epsilon$ does not seem realistic for practical applications. Indeed, $\mathcal{S}(\boldsymbol{X}_1^N)$ will almost surely vanish after a finite number of iterations if $\epsilon_k > e$ may happen. For that reason, we take here $e \geq \epsilon$. Assume that $\{\epsilon_k, \boldsymbol{x}_k\}$ are random variables over a fixed probability space, satisfying respectively (4) and (6), with $\mathcal{B}_k = \sigma(\{\epsilon_t, \boldsymbol{x}_t\}, t \leq k)$ the generated σ-algebra. Let $P(A)$ denote the probability of the event A. Following an approach similar to [6] we prove that the likelihood set converges to a fixed polytope centered at $\bar{\theta}$ assuming the following hypotheses:

H1: The sequence $\{\epsilon_k\}$ satisfies: $\exists C_1 > 0 \mid \forall k > 0\,,\ \forall j = 1, \dots, m\,,\ \forall u > 0$ small enough

$$\begin{aligned} P(\epsilon - \epsilon_k < u \mid \mathcal{B}_{k-1}, \boldsymbol{x}_k = \phi_j) &\geq C_1 u \ \text{ a.s.} \\ P(\epsilon + \epsilon_k < u \mid \mathcal{B}_{k-1}, \boldsymbol{x}_k = \phi_j) &\geq C_1 u \ \text{ a.s.} \end{aligned}$$

H2: The ϕ_j's span $\mathbb{R}^p$ and $\exists C_2 > 0 \mid \forall k > 0\,,\ \forall j = 1, \dots, m\,,\ P(\boldsymbol{x}_k = \phi_j \mid \mathcal{B}_{k-1}) \geq C_2$ a.s.

H3: The sequence $\{\epsilon_k, \boldsymbol{x}_k\}$ is asymptotically independent.

The theorem is then,

Theorem 1 *Under hypotheses H1 and H2, the likelihood set $\mathcal{S}(\boldsymbol{X}_1^N)$ converges to the polytope*

$$\bar{\mathcal{S}}(\boldsymbol{X}_1^\infty) = \{\theta \in \mathbb{R}^p \mid \\ -(e-\epsilon) \leq \phi_j^T(\theta - \bar{\theta}) \leq (e - \epsilon)\,,\ j = 1, \dots, m\}\,,$$

in the following sense:

$$\forall \theta \notin \bar{\mathcal{S}}(\boldsymbol{X}_1^\infty)\,,\ \exists \delta(\theta) > 0 \mid \\ \forall k > 0\,,\ P(|\boldsymbol{x}_k^T\theta - y_k| > e) > \delta(\theta)\,,$$

that is all θ not in $\bar{\mathcal{S}}(\boldsymbol{X}_1^\infty)$ is excluded from $\mathcal{S}(\boldsymbol{X}_1^k)$ with probability at least $\delta(\theta)$. If, moreover, H3 is satisfied, then

$$\forall \theta \notin \bar{\mathcal{S}}(\boldsymbol{X}_1^\infty)\,,\ \exists \text{ a.s. } N_0 \mid \forall N \geq N_0\,,\ \theta \notin \mathcal{S}(\boldsymbol{X}_1^N)\,.$$

3. BLIND BOUNDED-ERROR EQUALIZATION

Assume now that the true sequence of regressors $\{\boldsymbol{x}_k\}$ used to generate the data (2) is unknown, the issue being to recover it, as well as the parameters $\bar{\theta}$. We shall denote a guessed sequence of regressors by $\{\hat{\boldsymbol{x}}_k\}$, also assumed to belong to a finite set Φ as defined in (6). We say that a sequence $\hat{\boldsymbol{X}}_1^N$ is *consistent* with the data $\boldsymbol{y}_1^N$ if $\mathcal{S}(\hat{\boldsymbol{X}}_1^N) \neq \emptyset$. One wishes $\mathcal{S}(\hat{\boldsymbol{X}}_1^N)$ to be empty for N larger than some N_0 when $\hat{\boldsymbol{X}}_1^N$ does not coincide with $\boldsymbol{X}_1^N$. In fact, we shall see that this requirement is too strong in the general case, but that depending on the magnitude of the errors, only some particular sequences of regressors are consistent with the observations. For any $\hat{\boldsymbol{X}}_1^N$, define the sets

$$\begin{aligned} \mathcal{I}_1^N(j|i) &= \{k \in \{1, \dots, N\} \ / \ \hat{\boldsymbol{x}}_k = \phi_j, \boldsymbol{x}_k = \phi_i\}\,, \\ \Lambda_1^N(j) &= \cup_{i \in \{1, \dots, m\}} \mathcal{I}_1^N(j|i)\,, \end{aligned}$$

We shall require that the set Φ of regressors and the true value $\bar{\theta}$ of the parameters satisfy conditions of the following type:

H4: $\exists \gamma > 0 \ / \ \forall (i,j) \in \{1, \dots, m\}^2\,,$
$$i \neq j \ \Rightarrow \ |(\phi_i - \phi_j)^T \bar{\theta}| > \gamma\,.$$

Next theorem concerns the case where each observation can be associated with a unique regressor without error.

Theorem 2 *Assume that H4 is satisfied with $\gamma > 2(e + \epsilon)$ and that each ϕ_i appears at least once in $\boldsymbol{x}_1^N$, then*

$$\mathcal{S}(\hat{\boldsymbol{X}}_1^N) \neq \emptyset \Rightarrow \forall j \in \{1, \dots, m\}\,,\ \Lambda_1^N(j) = \mathcal{I}_1^N(j|\tau(j))\,,$$

with $\tau(.)$ a permutation of $\{1, \dots, m\}$.

This means that consistent guessed sequences correspond to permutations of the true sequence. It also shows that even in the presence of noise, the number of sequences $\{\hat{\boldsymbol{x}}_k\}$ consistent with the observations may remain bounded. Note that this bound may be quite large. Indeed, when the regressors are given by (3) with the alphabet (1) there might be $(2M)^p!$ consistent sequences. However, by exploiting the structure of the set (6) corresponding to the finite alphabet (1), we shall see in Section 4 that $\gamma > 2(e+\epsilon)$ in H4 is already sufficient to guarantee identifiability in the presence of bounded noise when the input sequence is rich enough.

Consider now the case of larger errors, with $2(e+\epsilon) > \gamma > 2e$. Next theorem states that a result similar to that of Theorem 2 will still hold almost surely for a finite N.

Theorem 3 *Assume that H4 is satisfied with $\gamma > 2e$, that exist $(i,j,l) \in \{1,\ldots,m\}^3$, $i \neq l$, such that $\exists k_1 \geq 1, k_2 \geq 2$ with $\boldsymbol{x}_{k_1} = \phi_i$, $\boldsymbol{x}_{k_2} = \phi_l$ and $\hat{\boldsymbol{x}}_{k_1} = \hat{\boldsymbol{x}}_{k_2} = \phi_j$, and that H1, H2, H3 are satisfied, then*

$$\exists a.s.\ N_0 \mid \forall N > N_0\,,\ \mathcal{S}(\hat{\boldsymbol{X}}_1^N) = \emptyset\,.$$

This means that all sequences not corresponding to permutations of the true sequence are almost surely not consistent after a finite number of observations.

Even in the rather favorable case $\gamma > 2e$, the average number of guessed consistent sequences for a given true sequence increases very fast with N (although much slower than the total number of possible sequences m^N), due to the fact that deterministic discrimination is impossible. However, next theorem presents a bound on the probability $P(N)$ that a randomly chosen sequence $\hat{\boldsymbol{X}}_1^N$ will be consistent with a random sequence $\boldsymbol{X}_1^N$ and shows that $m^N P(N)$ tends to zero exponentially fast as N increases when $\gamma > 2(e+\epsilon/3)$ when the ϵ_k's are uniformly distributed in $[-\epsilon,\epsilon]$, which complements the result in Theorem 3.

Theorem 4 *Assume that the sequences $\{\epsilon_k\}$ and $\{\boldsymbol{x}_k\}$ are independent, the distribution of the $\boldsymbol{x}_k$'s is uniform over Φ, and the distribution of the ϵ_k's is symmetric, that H4 is satisfied with $\gamma > 2e$, and that the $\hat{\boldsymbol{x}}_k$'s correspond to an i.i.d. sequence over Φ, independent of $\{\epsilon_k\}$ and $\{\boldsymbol{x}_k\}$, then*

$$m^N P(\mathcal{S}(\hat{\boldsymbol{X}}_1^N) \neq \emptyset) < \alpha(m)\lambda^N\,, \tag{7}$$

with $\alpha(m)$ not depending on N and λ depending only on the distribution of the ϵ_k's. When, moreover, ϵ_k is uniformly distributed in $[-\epsilon,\epsilon]$ with $\gamma > 2(e+\epsilon/3)$, then $\lambda < 1$ in (7).

After theses preliminary results, we now set the most important result of this communication. We show that the true input sequence will be recovered up to its sign in finite time in the presence of suitably bounded errors if the alphabet has the form (1).

4. IDENTIFIABILITY

Due to the form of the regressors (3) and alphabet (1), the sequence $\{\boldsymbol{x}_k\}$ lives on an oriented DeBruijn graph. The identifiability issue then consists in locating the process $\boldsymbol{x}_k$ on the graph (up to sign symmetry) when values $\boldsymbol{x}_k^T\bar{\theta}$ are observed.

Theorem 5 *Assume that all states are distinguishable (i.e. H4) and that the input sequence is such that at time K the process $\boldsymbol{x}_k$ has visited:*

- *2 loops, i.e. $\boldsymbol{x}_{k_i}^T\bar{\theta} = \boldsymbol{x}_{k_i+1}^T\bar{\theta}$, $i = 1,2$, such that $|\boldsymbol{x}_{k_1}^T\bar{\theta}| \neq |\boldsymbol{x}_{k_2}^T\bar{\theta}|$,*
- *all transitions on the shortest path between these two loops (these transitions do not need be visited consecutively)*

then the input sequence is identifiable at time K.

Note that under H2 the input sequence is a.s. identifiable in finite time and that H4 is much less conservative than the conditions used in [9, 1]. An important consequence of Theorem 5 is that identifiability can also be guaranteed in the presence of errors, provided H4 is satisfied with a suitable value of γ.

Corollary 1 *a/ Assume that H4 is satisfied with $\gamma > 2(e+\epsilon)$ and that the input sequence satisfies the conditions of Theorem 5, then the input sequence is identifiable at time K.*

b/ Assume that H4 is satisfied with $\gamma > 2e$, then H1, H2 and H3 imply that the input sequence is a.s. identifiable in finite time.

5. SIMULATION

We present here a simulation for a filter of lenght 2. The (unknown) channel parameters are $\theta_1 = -0.3, \theta_2 = 1.5$ (noted as a star * in the figures below). The noise has a raised-cosinus probability density fonction. The true bound error is $\epsilon = 0.05$ and the we take as known upper bound $e = 0.29$. In such case, H4 is verified.

Figures below illustrate our algorithm : The initialization of the polytope is drawn in dashed line on each figure. For the first step, we just examine two possibilities ($\begin{pmatrix}1\\1\end{pmatrix}$ and $\begin{pmatrix}1\\-1\end{pmatrix}$). The symmetric ones are not taking into account because identification is up to a sign. We then get figure 1 with two polytopes drawn in solid lines. Figure 2 represents the second step. Finally, figure 3 stands for the 4th step: the algorithm has then converged and the obtained polytope is the same as $\bar{\mathcal{S}}(\boldsymbol{X}_1^\infty)$.

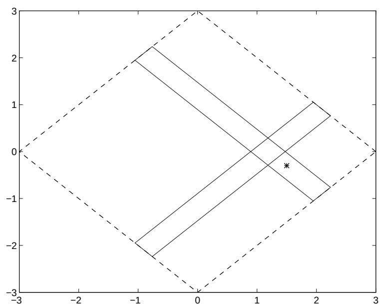

Figure 1: 1st step

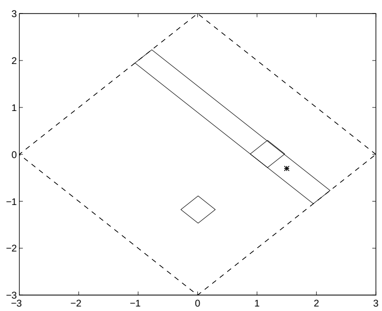

Figure 2: 2nd step

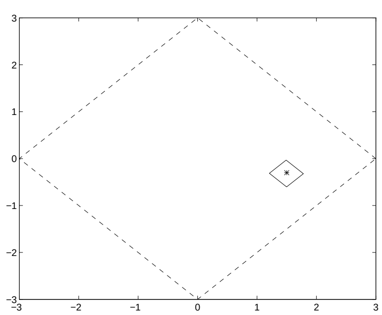

Figure 3: last step

6. CONCLUSIONS

Some properties of blind equalization in the context of bounded errors have been dicsussed. An interesting result is that identifiability can still be guaranteed in finite time provided that the bound on the error is not too large. The excitation conditions on the input sequence can be used to derive upper bounds on the expected waiting time for identifiability. Some simulations have been presented.

Further work will address the problem of joint identification of the bound on the errors and channel parameters through recursive L_∞ estimation, see [3].

7. REFERENCES

[1] F. Gustafsson & B. Wahlberg, "Blind equalization by direct examination of the input sequences", *IEEE Trans. Com.*, vol. 43, pp. 2213–22, 1995.

[2] Sessions SSAP1 and SSAP8 *Proceedings of the ICASSP*, Atlanta, Georgia, May 7-10, 1996.

[3] S. Icart, J. LeRoux, L. Pronzato, E. Thierry, A. Zhigljavski, "Blind Equalization in Presence of Bounded Errors with Unknown Bound", SPAWC , Paris, France, 1997.

[4] S.H. Mo & J.P. Norton, "Fast and robust algorithm to compute exact polytope parameter bounds", *Mathematics and Computers in Simulation*, vol. 32, pp. 481–93, 1990.

[5] J. Norton, editor, *Special Issue*, Int. J. Adapt. Control Signal Process., 1994, vol. 8.

[6] S. Veres & J.P. Norton, "Structure selection for bounded-parameter models: consistency conditions and selection criterion", *IEEE TAC*, vol. 36, pp. 474–481, 1991.

[7] E. Walter, editor, *Parameter Identification with Error Bounds*, Mathematics and Computers in Simulation, Dec. 1990, vol. 32(5,6).

[8] E. Walter & H. Piet-Lahanier, "Exact recursive polyhedral description of the feasible parameter set for bounded error models", *IEEE TAC*, vol. 34, pp. 911–15, 1989.

[9] D. Yellin & B. Porat, "Blind identification of FIR systems excited by discrete-alphabet inputs", *IEEE TSP*, vol. 41, pp. 1331–39, 1993.