

Sufficient Output Conditions for Identifiability in Blind Equalization

Dawei Huang and Fredrik Gustafsson

Abstract—The problem of input identifiability in blind deconvolution is considered where the input belongs to a known discrete alphabet. Input identifiability is an algorithm independent property, which does not necessarily imply channel identifiability. Sufficient conditions for input identifiability are derived in terms of algebraic relations on the observed output. It is shown how these new results relate to and unify other known sufficient conditions.

Index Terms—Blind equalization, deconvolution, equalizers, parameter estimation.

I. INTRODUCTION

THE typical application of blind deconvolution is in a digital communication system as depicted in Fig. 1. A signal a_n taken from a finite alphabet is distorted by a linear and time-invariant communication channel. The received signal x_n is processed in an equalizer and the output is denoted \hat{a}_n . The goal is to design the equalizer such that \hat{a}_n is as similar to a_n as possible. In the case of unknown signal a_n and channel, we have the problem of *blind equalization*.

The majority of the literature on blind equalization deals with the question of convergence of parameter values from a specific algorithm to those of the unknown channel inverse which is termed admissibility [1], [2], [5]. In contrast, there is limited work directed toward determining conditions on the channel output that permit the input data sequence to be identified often without recourse to assumptions on the statistics of the input, which is termed identifiability [4], [6], [7]. Identifiability is a fundamental algorithm independent concept that, for example, implies bounds on how fast practical blind algorithms can possibly converge. The conclusion of this paper is that the input can be determined as soon as we have found a linear combination $x_{n_1} \pm x_{n_2} \cdots \pm x_{n_m} = 0$ for any m . We will also relate this result to the sufficient conditions for identifiability given in [4] and [7].

II. PROBLEM FORMULATION AND NOTATION

Consider the following finite impulse response (FIR) channel model

$$x_n = \sum_{k=0}^{q-1} b_{k+1} a_{n-k}, \quad q \geq 2 \quad (1)$$

Paper approved by R. A. Kennedy, the Editor for Data Communications Modulation & Signal Design of the IEEE Communications Society. Manuscript received September 25, 1996; revised September 26, 1997.

D. Huang was with the Department of Electrical Engineering, Linköping University, S-581 83 Linköping, Sweden, on leave from the School of Mathematics, Queensland University of Technology, Brisbane, Q4001 Australia (e-mail: d.huang@fsc.qut.edu.au).

F. Gustafsson is with the Department of Electrical Engineering, Linköping University, S-581 83 Linköping, Sweden (e-mail: fredrik@isy.liu.se).

Publisher Item Identifier S 0090-6778(99)01922-4.

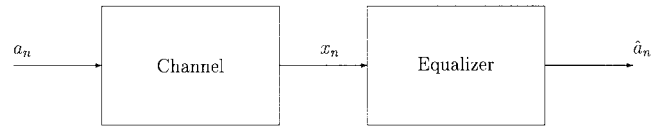


Fig. 1. A digital communication system.

where b_k are real valued channel parameters and the data a_k takes values in the alphabet $\{\pm 1, \pm 3, \dots, \pm(M-1)\}$, where M is a positive and even integer.

The vector (or sequence) $(a_1, a_2, \dots, a_n)'$ will be denoted a^n , and a_m^n denotes the vector $(a_m, a_{m+1}, \dots, a_n)'$, where $'$ means transpose. The vector of the q impulse response coefficients of the channel is denoted $b = (b_1, \dots, b_q)'$.

We can rewrite (1) in vector form as

$$x_q^n = \begin{bmatrix} a_q & a_{q-1} & \cdots & a_2 & a_1 \\ a_{q+1} & a_q & \cdots & a_3 & a_2 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ a_n & a_{n-1} & \cdots & \cdot & a_{n-q+1} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ b_q \end{bmatrix} = T_q(a^n) b \quad (2)$$

where $T_q(a^n)$ defines a Toeplitz matrix with q columns and elements a^n .

If the input sequence is known, the parameters can be computed by

$$b = (T_q(a^n)' T_q(a^n))^{-1} T_q(a^n)' x_q^n. \quad (3)$$

We define the following condition on the inverse to exist.

Definition 1: The input is said to be *persistently exciting* (PE) of order k at time n if $T_k(a^n)$ has full (column) rank.

That is, to determine the channel, the input must be PE of order q . This definition is quite common in adaptive control, see, e.g., [3], although a similar asymptotic condition is more common in system identification. However, the main purpose of this paper is to identify the input, and generally it is easier to estimate the input to a linear, time-invariant system than to identify the coefficients in the system.

Definition 2: The input is said to be *identifiable* from the observations at time n if the sequence a^n can be determined up to a constant c from the sequence x_q^n .

That the constant c is unavoidable is easily realized from (1). If ca_n defines another permissible input sequence, with elements in the given finite alphabet, then the channel $c^{-1}b$ gives exactly the same output. Typically, the sign of the input cannot be determined. Note that this definition is general and does not assume a specific algorithm used in blind equalization.

The sufficient condition we will present is split into one condition on the input sequence and one condition on the true channel parameters. The channel parameters are said to be *linearly independent* over the integer set $Z = (z_1, z_2, \dots)$ if

there does not exist a linear combination with coefficients in this set that equals zero, e.g., $z_1 b_3 + z_4 b_1 = 0$. The following four integer sets will be used:

$$Z_1 = \{0, \pm 2, \dots, \pm 2(M-1)\} \tag{4}$$

$$Z_2 = \{0, \pm 1, \dots, \pm 4^{q+1}(M-1)^{q+1}\} \tag{5}$$

$$Z_3 = \{0, \pm 1, \dots, \pm 8^{q+1}(M-1)^{q+1}\} \tag{6}$$

$$Z_{pe} = \left\{0, \pm 1, \dots, \pm 2(M-1)^{2q+1} q^{q/2} (n_0 - q + 1)^q\right\}. \tag{7}$$

In (7), n_0 denotes the first time instant when a^{n_0} is PE of order q .

III. SUFFICIENT CONDITIONS FOR IDENTIFIABILITY

This section presents the main theorem and points at some generalizations. Then a summary of known sufficient conditions for identifiability is given and the relation to our results is pointed out.

A. New Results

The following theorem says that if we get two consecutive observations of the same magnitude, then the input sequence can be determined up to a constant.

Theorem 1: The input sequence $a^n, n > q$, is identifiable if

- 1) the channel parameters b_1, \dots, b_q are linearly independent over Z_2 defined in (5).
- 2) For $\delta = \pm 1$, we have

$$x_n = \delta x_{n-1} \tag{8}$$

$$x_{n-1} \neq \delta x_{n-2}. \tag{9}$$

The proof is given in the Appendix. Here we remark that the conditions of the theorem can be satisfied with any degree of PE. This means that the input can be identifiable and still it may not be possible to determine the channel coefficients. Note that if condition (9) is not satisfied, both conditions (8) and (9) should be satisfied at an earlier time. Otherwise, we have observed the same signal all the time (perhaps with an alternating sign), which implies that the input sequence has been constant (or has had an alternating sign), see Lemma 1, and the definition of input identifiability holds.

With the same type of proof, we are able to show that it suffices that two observations are of the same magnitude at any time

$$\begin{aligned} x_n &= \delta x_m \\ x_{n-1} &\neq \delta x_{m-1}. \end{aligned} \tag{10}$$

If the integer set is extended to Z_3 , then we can go on and consider linear combinations of three observations

$$\begin{aligned} x_n + \delta_1 x_m + \delta_2 x_k &= 0 \\ x_{n-1} + \delta_1 x_{m-1} + \delta_2 x_{k-1} &\neq 0 \end{aligned} \tag{11}$$

where $\delta_i = \pm 1$. This can be further generalized to linear combination of m observations, $x_{n_1} + \delta_1 x_{n_2} + \dots + \delta_{m-1} x_{n_m} = 0$, and the integer sets need to be generalized to $Z_m = \{0, \pm 1, \dots, \pm 2^{m(q+1)}(M-1)^{q+1}\}$.

Finally, for complex input and channel response, these results are generalized by studying the real and imaginary part separately. Then it suffices that linear combinations as above of either the real or the imaginary part of x_n have appeared twice.

B. Relation to Known Results

The following sufficient condition is given in [7]:

- The input is identifiable if there exists four times n_1, n_2, n_3, n_4 with $|n_i - n_j| > q, \forall i, j$, where a first condition is

$$\begin{aligned} x_{n_1} &= x_{n_2} & x_{n_3} &= x_{n_4} \\ x_{n_1-1} &\neq x_{n_2-1} & x_{n_3-1} &\neq x_{n_4-1} \end{aligned} \tag{12}$$

and second condition is given for $x_{n_1+i} - x_{n_2+i}$ and $x_{n_3+i} - x_{n_4+i}$ for $i = 1, 2, \dots, q-1$.

The second condition and the algorithm are too complicated to be reviewed here. The first condition means that the observations can be divided into four nonoverlapping blocks, where the last observations are pairwise equal. We see that the output condition in (10) is strictly less restrictive. On the other hand, there is no condition on the channel in [7].

The result in [4] is as follows.

- *Necessary condition.* The input sequence and channel parameters are identifiable *only if* the input sequence is PE of order $2q-1$.
- *Sufficient condition.* The input sequence and channel parameters are identifiable *if* the input sequence is PE of order $2q-1$ and the true parameters are linearly independent over Z_{pe} as defined in (7).

Note that the condition of the signal concerns the input signal, which is not observable, in contrast to the new condition on the received signal.

In a sense, the sufficient condition in [4] is a limiting case of the conditions given here. As the number of observations in the linear combinations increases, the integer set increases. Note that $Z_1 \subset Z_2 \subset Z_3 \subset Z_{pe}$. That is, when Z_m is about as large as Z_{pe} , there must be a linear combination of the form (11) which is satisfied, if there is a solution at all.

C. An Example

Suppose that we have five observations from a binary channel of order $q=3$: $x_1 = 1, x_2 = -5, x_3 = -1, x_4 = 7, x_5 = 7$. We can assume that $a_5 = +1$. Then, since $x_4 = x_5$, we have from Lemma 1 $a_5 = a_4 = a_3 = a_2 \neq a_1$. Thus,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ -5 \\ -1 \\ 7 \\ 7 \end{bmatrix} = \begin{bmatrix} -1 & u_2 & u_1 \\ 1 & -1 & u_2 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix}.$$

Here $u_1 = a_0$ and $u_2 = a_{-1}$. There are five unknowns in four independent (bilinear) equations. However, we can find a

TABLE I

Combination (u_1, u_2)	(1,1)	(1,-1)	(-1,1)	(-1,-1)
Prediction Vector $H(u_1, u_2)$	$[\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, 1]$	$[\frac{1}{2}, 0, \frac{1}{2}, 1]$	$[\frac{1}{2}, \frac{1}{2}, 0, 1]$	$[\frac{1}{2}, 0, 0, 1]$
Prediction Error $H(u_1, u_2)X$	-6	0	-2	4

unique solution through the following steps. First, we have

$$X = \begin{bmatrix} x_1 - x_4 \\ x_2 - x_4 \\ x_3 - x_4 \\ x_4 \end{bmatrix} = \begin{bmatrix} -6 \\ -12 \\ -8 \\ 7 \end{bmatrix} = \begin{bmatrix} -2 & u_2 - 1 & u_1 - 1 \\ 0 & -2 & u_2 - 1 \\ 0 & 0 & -2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix}. \quad (13)$$

Compare with P defined in (15), and the expressions (17) and (18). Then, using (16) to compute the prediction vector $H(u_1, u_2)$, we get the results shown in Table I.

Thus, the combination corresponding to $H(u_1, u_2)X = 0$, i.e., $u_1 = 1$ and $u_2 = -1$, is the unique solution. Then it follows from (13) that

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} -2 & -2 & 0 \\ 0 & -2 & -2 \\ 0 & 0 & -2 \end{bmatrix}^{-1} \begin{bmatrix} -6 \\ -12 \\ -8 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}.$$

That is, in this case we can compute the true system

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ -5 \\ -1 \\ 7 \\ 7 \end{bmatrix} = \begin{bmatrix} -1 & -1 & 1 \\ 1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}.$$

Note that the first, rather conservative, condition in Theorem 1 is not satisfied, but we still get a unique solution.

APPENDIX PROOF OF THE THEOREM

A. A Preliminary Lemma

A key observation is given in the following lemma.

Lemma 1: Suppose b is linearly independent on $Z_1 = \{0, \pm 2, \dots, \pm 2(M-1)\}$. Then

$$\begin{aligned} x_n = x_{n+1} &\Leftrightarrow a_{n+1} = a_n = \dots = a_{n-q+1} = a, \\ x_n = -x_{n+1} &\Leftrightarrow (-1)^q a_{n+1} = (-1)^{q-1} a_n = \dots \\ &= a_{n-q+1} = a. \end{aligned}$$

Proof: Immediate by writing out $x_n \pm x_{n+1} = \sum_{k=0}^{q-1} b_{k+1}(a_{n-k} \pm a_{n-k-1})$. The coefficients of b_i belong to Z_1 , so by assumption the right-hand side is different from zero except when all coefficients are zero. \square

That is, if $x_{n-1} \pm x_n = 0$, we can conclude that the input has been constant (or alternating) for exactly $q+1$ samples.

B. Proof of Theorem 1

Consider the case $x_{n-1} = x_n$, the other case $x_{n-1} = -x_n$ is treated similarly. From Lemma 1 we have

$$\begin{bmatrix} x_{n-q-1} \\ x_{n-q} \\ \vdots \\ x_{n-2} \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} a_{n-q-1} & a_{n-q-2} & \dots & a_{n-2q+1} & a_{n-2q} \\ a & a_{n-q-1} & \dots & a_{n-2q+2} & a_{n-2q+1} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ a & a & \dots & a & a_{n-q-1} \\ a & a & \dots & a & a \\ a & a & \dots & a & a \end{bmatrix} b. \quad (14)$$

Let

$$P(u_1, \dots, u_q; u) = \begin{bmatrix} u_1 - u & u_2 - u & \dots & u_q - u \\ 0 & u_1 - u & \dots & u_{q-1} - u \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & u_1 - u \\ u & u & \dots & u \end{bmatrix} \quad (15)$$

be a $(q+1) \times q$ matrix. It is of rank q , so we can uniquely define a prediction vector $H(u_1, u_2, \dots, u_q; u) = [h_1, h_2, \dots, h_q, 1]$ by

$$H(u_1, u_2, \dots, u_q; u)P(u_1, u_2, \dots, u_q; u) = 0. \quad (16)$$

If we plug in the true input symbols in (16), then we can write

$$H(a_{n-q-1}, \dots, a_{n-2q}; a)(x_{n-q-1} - x_{n-1}) \quad (17)$$

$$= H(a_{n-q-1}, \dots, a_{n-2q}; a)P(a_{n-q-1}, \dots, a_{n-2q}; a)b = 0. \quad (18)$$

Now it can be shown that the elements in HP in (18) belong to Z_2 , so the assumption of linear independence implies that $HPb = 0$ is equivalent to $HP = 0$. The practical implication is as follows: if we have found an H satisfying (17), we have found the H making $HP = 0$ and thus also one possible solution to the blind equalization problem.

We know by assumption that there is a solution. We only need to show that the solution $HP = 0$ is unique. Suppose there is another matrix \bar{P} , so that $HP = H\bar{P} = 0$. Evaluating only the first column and using the structure of P and H gives $h_1(u_1 - u) + u = 0 = h_1(\bar{u}_1 - \bar{u}) + \bar{u}$, which (using $h_1 \neq 0$) implies $u = c\bar{u}$ and $u_1 = c\bar{u}_1$. By studying the second column we get $u_2 = c\bar{u}_2$, and so on. That is, the input sequence is uniquely identified up to a constant c , which is the definition of input identifiability.

REFERENCES

- [1] S. Bellini, "Blind equalization," *Alta Freq.*, vol. LVII, pp. 445–450, 1988.
- [2] A. Benveniste and M. Goursat, "Blind equalizers," *IEEE Trans. Commun.*, vol. COM-32, pp. 871–883, 1984.
- [3] G. C. Goodwin and K. S. Sin, *Adaptive Filtering, Prediction, and Control*. Englewood Cliffs, NJ: Prentice-Hall, 1984.
- [4] F. Gustafsson and B. Wahlberg, "Blind equalization by direct examination of the input sequences," *IEEE Trans. Commun.*, vol. 43, pp. 2213–2222, July 1995.
- [5] C. R. Johnson, "Admissibility in blind adaptive channel equalization," *IEEE Control Syst. Mag.*, vol. 11, pp. 3–15, 1991.
- [6] H. Liu, G. Xu, and L. Tong, "A deterministic approach to blind equalization," in *Conf. Rec. 27th Asilomar Conf. Signals, Systems, Computers*, 1993, pp. 751–755.
- [7] D. Yellin and B. Porat, "Blind identification of FIR systems excited by discrete-alphabet inputs," *IEEE Trans. Signal Processing*, pp. 1331–1339, 1993.