SEMI-BLIND IDENTIFICATION OF FINITE IMPULSE RESPONSE CHANNELS

Jonathan H. Manton, Yingbo Hua, Yufan Zheng and Cishen Zhang

Department of Electrical and Electronic Engineering,
University of Melbourne, Parkville, Victoria 3052, Australia.
j.manton@ee.mu.oz.au

ABSTRACT

It is a standard result that a finite impulse response channel of length L can be uniquely identified by feeding in a known (and persistently exciting) sequence of $2L-1$ consecutive data points. Equivalently, given only $2L-2$ consecutive data points, the channel can be uniquely identified up to a multiplicative constant. This paper significantly extends the identifiability criterion to the case when the known inputs are non-consecutively located. It is argued that by introducing $2L-1$ non-consecutively spaced zeros into the input stream, for almost all input sequences, the channel can be uniquely identified up to a multiplicative constant. Furthermore, the result can be extended to the case when the known inputs are non-zero, in which case the channel can almost always be identified uniquely. To arrive at these results, general properties of systems of polynomial equations are derived. These properties do not seem to have appeared in the literature before.

Key words: Algebraic geometry, Commutative algebra, Polynomial equations, Semi-blind identification, Finite impulse response channels.

1. INTRODUCTION

The identification of an FIR (finite impulse response) channel is a fundamental problem in signal processing. Identification of the impulse response of the channel is typically based on the knowledge of the output of the channel, the length of the channel, and on some known property of the input to the channel. The type of identification depends upon which property of the input signal is known. A review of blind identification techniques is given in [1], while a fast algorithm for blind identification is given in [5].

Recently, in [4], the input sequence was broken up into fixed length blocks, and a zero appended to each block. These added zeros allowed the input signal to be treated in a stochastic framework. (It was shown that the input was cyclostationary.) It was proved that asymptotically, as the number of blocks goes to infinity, the channel can be identified up to a multiplicative constant. The present paper considers the same problem in a deterministic framework, and shows that only a finite number of blocks are required to uniquely identify the channel up to a multiplicative constant.

In the deterministic framework, the semi-blind identification problem corresponds to solving a system of polynomial equations. The relevant areas of mathematics which extensively study polynomial equations are commutative algebra and algebraic geometry.

Jonathan Manton acknowledges the support of the Australian Telecommunication and Engineering Research Board and the Cooperative Research Centre for Sensor Signal and Information Processing.

From a practical point of view, solving systems of polynomial equations has been extensively studied; see for example "elimination theory" in [3]. However, the main theoretical result, Theorem 1, has not (to our knowledge) appeared in the literature previously, presumably due to the fact that the specific problem has not been of great relevance to mathematicians.

Section 2 describes the semi-blind identification problem and gives an example. After background material is presented in Section 3, the main results of this paper are given in Section 4. Further work is briefly mentioned in Section 5.

2. SEMI-BLIND IDENTIFICATION

The aim of semi-blind identification is to determine the finite impulse response (FIR) channel based on the channel output and partial knowledge of the channel input. For ease of presentation, it is assumed that every Tth input ($T \geq 2$) is zero. These zero inputs are referred to as non-consecutive training data.

Denote the input sequence by $\cdots, x_{-1}, x_0, x_1, \cdots$; the output sequence by $y_1, y_2, \cdots$; and the channel by $h_0, h_1, \cdots, h_{L-1}$, where L is the length of the channel. The input, output and channel coefficients are all sequences of complex numbers. For mathematical convenience, it is assumed that $h_0 = 1$. (With zero valued training data, the channel can only be identified up to a multiplicative constant. Therefore, the value of h_0 (which is assumed non-zero) can be arbitrarily chosen.)

The input and output are related by the N equations:

$$y_n = \sum_{k=0}^{L-1} h_k x_{n-k}, \quad n = 1, \cdots, N \tag{1}$$

The training data equations are:

$$x_T = x_{2T} = x_{3T} = \cdots = 0 \tag{2}$$

By choosing $N = (2L-2)T$, (1) and (2) correspond to $N + (2L-2)$ equations in $N + (2L-2)$ unknowns, and by choosing $N = (2L-1)T$, (1) and (2) correspond to $N + (2L-1)$ equations in $N + (2L-2)$ unknowns. In the former case, there are n equations in n unknowns, and in the latter case, with one extra training data, there are $n+1$ equations in n unknowns. Furthermore, these equations are all polynomials.

Example: Consider the simple case $L = 2$ and $T = 3$. The channel equations (1) become, after substituting the training data (2) (i.e., $x_3 = x_6 = 0$):

$$\begin{aligned} y_1 &= x_1 + hx_0 & y_2 &= x_2 + hx_1 & y_3 &= hx_2 \\ y_4 &= x_4 & y_5 &= x_5 + hx_4 & y_6 &= hx_5 \end{aligned} \tag{3}$$

There are 6 equations and 6 unknowns ($h, x_0, x_1, x_2, x_4, x_5$). By eliminating x_4, x_5 in (3), the polynomial $y_4h^2 - y_5h + y_6 = 0$ results, showing that there are in general two different values of h.

Three important points can be made. 1) Unlike the linear case when n equations in n unknowns uniquely determine a point, in the polynomial case, a finite number of solutions are possible. 2) For some exceptional values of the input, there might be an infinite number of solutions. In the example above, if $y_4 = y_5 = 0$, h can be arbitrary. Therefore, any general theory must take into account these exceptions. 3) The equations in (3) are bilinear.

If three more outputs are observed, i.e., $y_7 = x_7,\ y_8 = x_8 + hx_7,\ y_9 = hx_8$, then a second quadratic in h can be obtained, namely $y_7h^2 - y_8h + y_9 = 0$.

Note that if the inputs are chosen at random, then the outputs are random too. In particular then, the equations $y_4h^2 - y_5h + y_6 = 0$ and $y_7h^2 - y_8h + y_9 = 0$ are "random", and so intuitively, with probability one, they will have only one solution (remember that the output is such that at least one solution must exist).

Summary of key results: The example given above illustrates the main results this paper attempts to prove. Assume that the input sequence is random. By using $2L - 2$ training data, there are n polynomial equations in n variables, and hence, with probability one, a finite number of solutions. By using $2L - 1$ training data, there are $n + 1$ polynomial equations in n variables, and hence, with probability one, a unique solution.

3. REVIEW OF BACKGROUND MATERIAL

This section briefly outlines the notation and results required from algebraic geometry and commutative algebra. The book [3] is recommended as an easy yet comprehensive introduction to algebraic geometry. Other references include [6, 8]. For commutative algebra, the book [2] is recommended.

This paper uses the word **variety** to mean an affine variety over the algebraically closed field $\mathcal{C}^n$, for some positive integer n. A set V of points in $\mathcal{C}^n$ is thus a variety if it corresponds to the set of all solutions of a finite system of polynomials. To be precise, define the set

$$\mathbf{V}(f_1, \cdots, f_r) = \{(x_1, \cdots, x_n) \in \mathcal{C}^n : f_1(x_1, \cdots, x_n) = 0, \cdots, f_r(x_1, \cdots, x_n) = 0\} \quad (4)$$

to be the set of all solutions of the r polynomials $f_1, \cdots, f_r \in \mathcal{C}[x_1, \cdots, x_n]$, where the notation $\mathcal{C}[x_1, \cdots, x_n]$ represents the polynomial ring consisting of all polynomials with complex coefficients in the n variables $x_1, \cdots, x_n$. Then a set $V \subset \mathcal{C}^n$ is a variety if it can be written as $V = \mathbf{V}(f_1, \cdots, f_r)$ for some r polynomials $f_1, \cdots, f_r \in \mathcal{C}[x_1, \cdots, x_n]$.

The **Zariski topology** is the topology induced on $\mathcal{C}^n$ by defining a set to be closed iff it is a variety. The word **dense** is taken with respect to the Zariski topology, i.e., a set is dense in V iff the smallest variety containing the set is V. The **closure** of a set E (denoted $\overline{E}$) is therefore the smallest variety containing E.

The union of a finite number of varieties is again a variety. The intersection of an arbitrary number of varieties is again a variety. A variety is **irreducible** if it cannot be split into two smaller varieties, i.e., $V \subset \mathcal{C}^n$ is irreducible if $V = V_1 \cup V_2$ implies either $V_1 = V$ or $V_2 = V$.

A variety $V \subset \mathcal{C}^n$ can always be **decomposed** into a finite number of **irreducible components**, i.e., $V = V_1 \cup \cdots \cup V_k$ where $V_1, \cdots, V_k$ are irreducible varieties, and $V_i \not\subset V_j$ for $i \neq j$. (The decomposition is unique up to the order in which $V_1, \cdots, V_k$ are written.)

The **dimension** of a variety can be defined in a number of equivalent ways. A precise definition is outside the scope of this paper; see [3].

Proposition 1 *Let W be a subvariety of an irreducible variety V (i.e., $W \subset V$). Then $\dim W < \dim V$ iff W is a proper subvariety (i.e., $W \neq V$).*

Remark: If the variety V is reducible, Proposition 1 can be used in conjunction with the fact that if $V = V_1 \cup V_2$, then $\dim V = \max(\dim V_1, \dim V_2)$.

4. MAIN RESULTS

In Section 2 it was shown that in the semi-blind identification problem, the channel is identified by solving a system of polynomial equations. If these equations have a unique solution, then the channel is uniquely identified. Therefore, this section primarily studies the number of solutions of a system of polynomial equations.

A property can hold "almost everywhere" rather than everywhere. Definition 1 below gives an appropriate definition of "almost everywhere" based on the dimension of a variety. To visualise the definition, recall that a point, a line, and a plane have dimension 0, 1 and 2 respectively.

Definition 1 *A property will be said to hold for* **almost all** *points of a variety V if the set of points for which the property does not hold (called the* **exceptional points***) is contained in a subvariety W of V with a lower dimension, i.e., if $\dim W < \dim V$. When the variety V is understood from the context, the property will be more succintly said to hold* **almost everywhere**.

Remark: Proposition 1 provides a convenient criterion for determining if a property holds almost everywhere.

The definition given above is consistent with the probabilistic (measure-theoretic) notion of "with probability one". Lemma 1 below shows that if a point is chosen at random, with probability one, it will not be an exceptional point.

Lemma 1 *Let $[x_1, \cdots, x_m]'$ be a vector of real random variables. Assume each random variable is absolutely continuous and is independent. Then for any real vector $[y_1, \cdots, y_m]'$ (possibly depending on $[x_1, \cdots, x_m]'$), the point $\mathbf{z} = [x_1 + jy_1, \cdots, x_m + jy_m]' \in \mathcal{C}^m$ lies in a proper subvariety of $\mathcal{C}^m$ with probability zero.*

Remark: The assumption of independence can be relaxed. It suffices that the probability density of each random variable conditioned on the others is absolutely continuous. For example, $\mathbf{z}$ can be a Gaussian random vector with arbitrary (but non-singular) correlation.

Proposition 2 below highlights key properties of polynomial mappings. Note that the notion of almost everywhere is preserved by inverse polynomial mappings.

Proposition 2 *The inverse image of a closed set under a polynomial mapping is closed. Furthermore, the inverse mapping is order preserving. Specifically, for any $f_1, \cdots, f_r \in \mathcal{C}[x_1, \cdots, x_n]$,*

define the polynomial mapping F from $\mathcal{C}^n$ to $\mathcal{C}^r$ by

$$F(x_1, \cdots, x_n) = (f_1(x_1, \cdots, x_n), \cdots, f_r(x_1, \cdots, x_n)) \quad (5)$$

For any set $V \subset \mathcal{C}^n$, define the image of V as

$$F(V) = \{(c_1, \cdots, c_r) \in \mathcal{C}^r : (c_1, \cdots, c_r) = F(x_1, \cdots, x_n), (x_1, \cdots, x_n) \in V\} \quad (6)$$

For any set $W \subset \mathcal{C}^r$, define the inverse image of W as

$$F^{-1}(W) = \{(x_1, \cdots, x_n) \in \mathcal{C}^n : F(x_1, \cdots, x_n) \in W\} \quad (7)$$

Then 1) for any variety $W \subset \mathcal{C}^r$, $F^{-1}(W)$ is a variety; 2) for any varieties $W' \subsetneq W \subset \mathcal{C}^r$, $F^{-1}(W') \subset F^{-1}(W)$; and 3) for any variety $V \subset \mathcal{C}^n$, if $W \subsetneq \overline{F(V)}$, then $F^{-1}(W) \subsetneq V$.

Before the main theorems concerning systems of polynomial equations are given, the term "algebraic dependence" is defined.

Definition 2 *The r polynomials $f_1, \cdots, f_r \in \mathcal{C}[x_1, \cdots, x_n]$ are said to be* **algebraically dependent** *if there exists a non-zero polynomial $g \in \mathcal{C}[f_1, \cdots, f_r]$ such that $\forall (x_1, \cdots, x_n) \in \mathcal{C}^n$,*

$$g(f_1(x_1, \cdots, x_n), \cdots, f_r(x_1, \cdots, x_n)) = 0 \quad (8)$$

(This is succintly written as $g(f_1, \cdots, f_r) \equiv 0$.) Otherwise, the equations are said to be **algebraically independent**.

Remark: It is straightforward to prove that for any complex constants c_i, $i = 1, \cdots, r$, the polynomials $f_i - c_i$ are algebraically independent iff the f_i are themselves algebraically independent. Also, any set of $n+1$ polynomial equations in $\mathcal{C}[x_1, \cdots, x_n]$ (i.e., in n variables) are dependent.

Theorem 1 gives a general statement concerning the number of solutions of a set of n polynomial equations in n unknowns.

Theorem 1 *Given n polynomials $f_1, \cdots, f_n \in \mathcal{C}[x_1, \cdots, x_n]$, consider the system of equations $f_i(x_1, \cdots, x_n) = c_i$ for arbitrary complex constants c_i ($i = 1, \cdots, n$). Let $C \subset \mathcal{C}^n$ denote the image of the f_i, i.e., $C = F(\mathcal{C}^n)$ (c.f., (6)). Then:*

1. *$\overline{C}$ is an irreducible variety, where $\overline{C}$ denotes the Zariski closure.*
2. *The f_i are algebraically independent iff $\overline{C} = \mathcal{C}^n$.*
3. *If the equations f_i are algebraically independent, then for almost all $(c_1, \cdots, c_n) \in \overline{C}$, $f_i = c_i$, $i = 1, \cdots, n$ has a finite number of solutions.*
4. *If the equations f_i are algebraically dependent, then for almost all $(c_1, \cdots, c_n) \in \overline{C}$ (in fact for all $(c_1, \cdots, c_n) \in C$), $f_i = c_i$, $i = 1, \cdots, n$ has an infinite number of solutions.*

Remark 1: While Theorem 1 should not be a surprising result to an expert in commutative algebra, we have been unable to find any statement of it in the literature. Furthermore, it is not straightforward to prove Theorem 1.

Remark 2: The example below shows how n independent equations can define a variety having an infinite number of solutions. Let $f_1(x_1, x_2) = x_1(x_1 - x_2)$ and $f_2(x_1, x_2) = x_2(x_1 - x_2)$. Then f_1 and f_2 are independent. [The Jacobian (see below) is $|J(f_1, f_2)| = 2(x_1 - x_2)^2 \not\equiv 0$ and so, by Theorem 2, f_1 and f_2 are independent.] However, $\mathbf{V}(f_1, f_2) = \mathbf{V}(x_1 - x_2)$ and thus has an infinite number of solutions.

By applying Proposition 2 to Theorem 1, Lemma 2 results.

Lemma 2 *Consider the n polynomial equations $f_i = c_i$ in the n complex variables $x_1, \cdots, x_n$ for arbitrary complex constants c_i. The space $\mathcal{C}^n$ can be partitioned into two disjoint sets, X_f and X_i. A point $(p_1, \cdots, p_n) \in \mathcal{C}^n$ lies in X_f if there are a finite number of solutions $(x_1, \cdots, x_n)$ to the n equations $f_i(x_1, \cdots, x_n) = f_i(p_1, \cdots, p_n)$, $i = 1, \cdots, n$. Otherwise, the point $(p_1, \cdots, p_n)$ lies in X_i (i.e., infinite number of solutions). Then almost all points of $\mathcal{C}^n$ lie in the set X_f if and only if the f_i are algebraically independent.*

Before Lemma 2 can be applied to the semi-blind identification problem, a test for determining if a set of equations are independent is required. Combining ideas from both analytic and algebraic geometry, the criterion given in Theorem 2 was devised. It is based upon the Jacobian matrix, which is now defined:

Definition 3 *Given r polynomials $f_1, \cdots, f_r \in \mathcal{C}[x_1, \cdots, x_n]$, define the* **Jacobian matrix** *to be the $r \times n$ matrix of partial derivatives:*

$$J(f_1, \cdots, f_r) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_r}{\partial x_1} & \cdots & \frac{\partial f_r}{\partial x_n} \end{bmatrix} \quad (9)$$

Evaluating this matrix at $p \in \mathcal{C}^n$ gives a matrix of numbers denoted by $J_p(f_1, \cdots, f_r)$. The **Jacobian** *is the determinant of the Jacobian matrix, written $|J|$. (Clearly, the Jacobian is only defined if the Jacobian matrix is square, i.e., $r = n$.) Note that the Jacobian is a polynomial in $\mathcal{C}[x_1, \cdots, x_n]$.*

Theorem 2 *The n polynomials $f_i(x_1, \cdots, x_n)$, $i = 1, \cdots, n$ in the n complex variables $x_1, \cdots, x_n$ are algebraically independent if the $n \times n$ Jacobian matrix $J(f_1, \cdots, f_n)$ (defined in (9)) is not everywhere singular, i.e., if there exists a point $p \in \mathcal{C}^n$ such that $|J_p(f_1, \cdots, f_n)| \neq 0$.*

At this point, the following theorem can be given, ensuring that for $2L - 2$ known inputs, the channel (of length at most L) can be identified up to at most a finite number of possibilities for almost all input sequences. Applying Lemma 1, this says that if the input is chosen at random, then with probability one the channel can be identified up to a finite number of possibilities.

Theorem 3 *By choosing $N = (2L-2)T$, the semi-blind identification equations (1) and (2) will have a finite number of solutions for almost all combinations of input sequences and channels.*

PROOF. The N equations (1), after substituting in (2), have the following Jacobian matrix ($L = 2$, $T = 3$ case presented for concreteness):

$$J(f_1, \cdots, f_6) = \begin{bmatrix} h & 1 & 0 & 0 & 0 & x_0 \\ 0 & h & 1 & 0 & 0 & x_1 \\ 0 & 0 & h & 0 & 0 & x_2 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & h & 1 & x_4 \\ 0 & 0 & 0 & 0 & h & x_5 \end{bmatrix} \quad (10)$$

Since there exists a value of $(h, x_0, x_1, x_2, x_4, x_5)$ for which J is non-singular, by Theorem 2, the N equations are algebraically independent. By Lemma 2, for almost all values of the channel and

input $(h, x_0, x_1, x_2, x_4, x_5)$, there are a finite number of solutions. □

The following conjecture is also given. It is based on the following intuitive argument[1]: We conjecture that a set of polynomials in $\mathcal{C}[x_1, \cdots, x_n]$ has a unique solution almost everywhere if and only if the field generated by the polynomials is the same as the field $\mathcal{C}(x_1, \cdots, x_n)$. Furthermore, given n independent equations, they generate a field, say F, with the same transcendence degree as $\mathcal{C}(x_1, \cdots, x_n)$, namely n. Therefore, the field $\mathcal{C}(x_1, \cdots, x_n)$ is a finite extension of F. It is well known that in this case, only a single polynomial f need be added to F to obtain $\mathcal{C}(x_1, \cdots, x_n)$. Furthermore, and roughly speaking, almost any f will do.

Conjecture 1 *By choosing $N = (2L-1)T$, the semi-blind identification equations (1) and (2) will have a unique solution for almost all combinations of input sequences and channels.*

5. DISCUSSION OF FUTURE WORK

For reasons of space and clarity, this paper has only considered the case when the training data is zero and is periodically spaced. It should be clear that the actual training data used is unimportant; the channel can still be identified with probability one. Furthermore, with non-zero training data, it is possible to identify the channel completely, rather than up to a multiplicative constant.

Symbolic mathematics packages provide routines for solving systems of polynomial equations. Elimination theory can be applied to specific cases to determine what sequences of inputs cannot be identified etc.

In terms of actually finding the channel, we remark that due to ill-conditioning, it is not recommended that the semi-blind identification problem be solved by elimination. A recent paper [7] presents a more stable technique for solving a system of polynomial equations.

Finally, it is hoped that we are able to rigorously prove Conjecture 1 in the near future.

6. CONCLUSION

This paper has three main contributions. Firstly, it has placed the semi-blind identification problem in an algebraic framework. This allows existing knowledge of polynomial equations found in the mathematics literature to be applied to the engineering problem of semi-blind identification. Secondly, non-trivial theorems concerning the theoretical nature of systems of polynomial equations have been derived (most significantly Theorem 1). Thirdly, by applying the theorems to the semi-blind identification problem, it was shown that an FIR channel can be semi-blindly identified using only a finite number of training data. As outlined in Section 5, the methods used in this paper can be used to answer other questions about semi-blind identification.

7. REFERENCES

[1] K. Abed-Meraim, W. Qiu, and Y. Hua. Blind system identification. *Proceedings of the IEEE*, 85(8):1310–1322, August 1997.

[2] M. F. Atiyah and I. G. MacDonald. *Introduction to Commutative Algebra*. Addison-Wesley, 1969.

[3] D. A. Cox, J. B. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Springer-Verlag, 2nd edition, 1996.

[4] G. B. Giannakis. Filterbanks for blind channel identification and equalization. *IEEE Signal Processing Letters*, 4(6):184–187, June 1997.

[5] Y. Hua. Fast maximum likelihood for blind identification of multiple FIR channels. *IEEE Transactions on Signal Processing*, 44(3):661–672, March 1996.

[6] K. Kendig. *Elementary Algebraic Geometry*. Graduate Texts in Mathematics 44. Springer-Verlag, 1977.

[7] D. Manocha. Solving systems of polynomial equations. *IEEE Computer Graphics and Applications*, 14:46–55, March 1994.

[8] I. R. Shafarevich. *Basic Algebraic Geometry*. Springer Study Edition. Springer-Verlag, 1977.

[1] The argument was outlined to us by Professor John Groves of the mathematics department, The University of Melbourne.