A Deterministic Approach to Blind Symbol Estimation

Hui Liu and Guanghan Xu

Dept. of Electrical & Computer Engineering
The University of Texas at Austin
Austin, TX 78712-1084

Abstract

A blind symbol estimation technique for digital communications is developed by exploiting special data structure of the oversampled channel output. The proposed method achieves direct symbol estimation without determining the channel characteristics. Moreover, if the transmitting symbols belong to a finite set of alphabets, the new approach can be extended to handle multiple co-channel sources.

1 Introduction

The problem of blind channel identification using the second-order statistics of the *oversampled channel output* has drawn considerable attention recently following the pioneering work of Tong, Xu and Kailath [8]. Many interesting results have been developed [9, 2, 1, 4, 6, 5]. Although the second-order approach has significant advantages over the conventional higher order methods which may not be data efficient, it is still statistics-based and may suffer from finite data effect when dealing with an extremely short sample sequence. As an attempt to remedy this problem and provide an algebraic interpretation of the blind identification problem, we establish a deterministic model for the multichannel communication system by treating the input sequence as an unknown deterministic signals [3]. Similar data models have been adopted by various researchers and many data efficient parametric blind estimation approaches have been developed [6, 4, 5, 3].

The focus of the aforementioned methods has been on the blind identification of the channel. In this paper, we extend our investigation and propose a new algorithm which accomplishes symbol estimation directly from the oversampled output data of the channel.

2 Problem Statement

Using the spatial and/or temporal oversampling techniques [8], a communication channel can be cast into a single-input multiple-output (SIMO) system with multiple subchannels. Denote $x_i(\cdot)$ the outputs from the ith subchannel, they are related to the FIR channel response $\{h_i(\cdot)\}$ and the system inputs $s(\cdot)$ by

$$x_i(k) = \sum_{j=0}^{L} h_i(j)s(k-j), \qquad i = 1 \cdots M, \tag{1}$$

where M is the number of subchannels and L the maximum order of all subchannels. The case of principal interest herein is the estimation of the input symbols $s(\cdot)$ from *finite* data samples without the knowledge of the channel characteristics.

Let $\mathbf{x}_k = [x_1(k) \cdots x_M(k)]^T$, $\mathbf{h}_l = [h_1(l) \cdots h_M(l)]^T$. For a finite number of data samples ($k = L+1, \cdots, N$), (1) can be rewritten in a matrix form:

$$\underbrace{[\mathbf{x}_{L+1} \ \cdots \ \mathbf{x}_N]}_{\mathbf{X}} = \underbrace{[\mathbf{h}_L \ \cdots \mathbf{h}_0]}_{\mathbf{H}} \times \underbrace{\begin{bmatrix} s(1) & \cdots & s(N-L) \\ \vdots & \cdots & \vdots \\ s(L+1) & \cdots & s(N) \end{bmatrix}}_{\mathbf{S}}. \tag{2}$$

The subspace defined by the rows of $\mathbf{S}$ is called the *signal subspace* and is denoted as $\mathbf{V}_s$. Its orthogonal complement $\mathbf{V}_o$ is referred to as the *null subspace*. When $\mathbf{H}$ is of full column rank, $\mathbf{X}$ has the same row subspace as $\mathbf{V}_s$.

Since the channel matrix $\mathbf{H}$ has a dimension of $M \times (L+1)$, it can not be of full column rank if $M < L+1$. In such a case, we need to smooth the output data vectors so that the rank of the channel matrix can be restored (see [8] for details).

1058-6393/95 $4.00 © 1995 IEEE

$$\underbrace{\begin{bmatrix} \mathbf{x}_{L+1} & \cdots & \mathbf{x}_{N-K+1} \\ \vdots & \cdots & \vdots \\ \mathbf{x}_{L+K} & \cdots & \mathbf{x}_N \end{bmatrix}}_{\mathbf{X}(K)} = \underbrace{\begin{bmatrix} \mathbf{h}_L & \cdots & \mathbf{h}_0 & \cdots & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{h}_L & \cdots & \mathbf{h}_0 \end{bmatrix}}_{\mathbf{H}(K),\ r\ \text{blocks}} \times \underbrace{\begin{bmatrix} s(1) & s(2) & \cdots & s(N-r+1) \\ s(2) & s(3) & \cdots & s(N-r+2) \\ \vdots & \vdots & \cdots & \vdots \\ s(r) & s(r+1) & \cdots & s(N) \end{bmatrix}}_{\mathbf{S}(r)} \quad (3)$$

where $r = L + K$ and K is the smoothing factor. As K increases, $\mathbf{H}(K)$ will eventually have more rows than columns. The row subspace $\mathbf{X}(K)$ becomes the signal subspace if $\mathbf{H}(K)$ is of full column rank.

In blind estimation problems, neither $\mathbf{S}$ nor $\mathbf{H}$ in (2) is known. A further investigation of (3) shows that they both contain rich structure (Hankel) information which can possibly be exploited for blind estimation. In the following sections, we provide an alternate conceptual framework for understanding the deterministic blind symbol estimation problem and demonstrate the feasibility of estimating the inputs directly from the row subspace of $\mathbf{X}$.

3 Blind Symbol Estimation

First, we present an important theorem which lays the groundwork for the development of our Blind Symbol Estimation (BSE) algorithm. We define $\mathbf{S}(r)$ in (3) the *finite* input matrix of order r. If $r = 1$, $\mathbf{S}(1)$ becomes a row vector and we refer $\mathbf{s} = \mathbf{S}(1)^H$ as the input vector.

Theorem 1 *The input vector* $\mathbf{s}$ *can be uniquely determined up to a scalar from the row subspace of* $\mathbf{S}(r)$ *if* $\mathbf{s}$ *has a sufficient number of modes*[1].

Proof:
See Appendix A. □

Since the row subspace of the input matrix $\mathbf{S}(r)$ is often shared by the output data matrix $\mathbf{X}(K)$, the physical significance of Theorem 1 lies in the fact that it assures the identifiability of the input sequences from the row subspace of the data matrix!

[1] See [3] for the definition of a mode. It can be interpreted an embedded exponential in the input sequence.

3.1 The Proposed Method

Following the assertion in Theorem 1, we derive the BSE method as follows.

Assuming that $\mathbf{H}(K)$ is of full column rank, the signal and null subspaces of of $\mathbf{S}(r)$ can be calculated from the output data matrix $\mathbf{X}(K)$. By a property of Hankel matrices, it is shown in [10] that given the null subspace of $\mathbf{S}(r)$, the null subspace of $\mathbf{S}(r-1)$ can be constructed as $\begin{bmatrix} \mathbf{V}_o(r) & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_o(r) \end{bmatrix}$, where $\mathbf{0}$ is a $(N - 2r + 1) \times 1$ vector. Following this procedure, we repeat the construction and remove the redundant rows r times, the null subspace of $\mathbf{S}(1)$ has the following form,

$$\mathbf{V} = \underbrace{\begin{bmatrix} \mathbf{V}_o(r) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_o(r) & \mathbf{0} & \vdots \\ \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{V}_o(r) \end{bmatrix}}_{r\ \text{blocks}}. \quad (4)$$

It is readily seen that the input vector $\mathbf{s}$ is the unique non-trivial solution of the overdetermined linear equations $\mathbf{V}\mathbf{s} = \mathbf{0}$.

In practical situations, only noise corrupted data is available. The inputs can then be estimated by finding the least square solution. The whole algorithm can be summarized as

1. calculate the null space of $\mathbf{S}$ from $\mathbf{X}$;
2. construct the $\mathbf{V}$ matrix as in (4);
3. estimate $\mathbf{s}$ by solving $\mathbf{V}\mathbf{s} = \mathbf{0}$.

The strength of the proposed algorithm is its effectiveness in dealing short data sequences, which makes it particularlly suitable for wireless systems where the environment changes rapidly. For a large size of data samples, statistics-based methods with better asymptotic efficiency are preferred.

4 Extension to Multiple Sources

For P ($P > 1$) sources, (2) and (3) become $\mathbf{X} = \sum_{i=1}^{P} \mathbf{H}_i \mathbf{S}_i = [\mathbf{H}_1 \ \cdots \ \mathbf{H}_P] [\mathbf{S}_1^T \ \cdots \ \mathbf{S}_P^T]^T$. The row subspace of $[\mathbf{S}_1^T \cdots \mathbf{S}_P^T]^T$ can still be computered from $\mathbf{X}$ provided that the channel matrix $[\mathbf{H}_1 \ \mathbf{H}_2 \ \cdots \ \mathbf{H}_P]$ is of full column rank (To achieve this, smoothing of the output vectors may be required).

Theorem 2 *Let* $\mathbf{s}_i$, $i = 1, \cdots, P$ *be multiple input vectors with its corresponding Hankel matrices* $\{\mathbf{S}_i(r)\}$ *and* $\mathbf{t}$ *be another vector with its own Hankel matrix* $\mathbf{T}(r)$ *that is defined as* $\mathbf{S}(r)$ *in (3). If*

1. $\mathcal{R}\{\mathbf{T}(r)\} \subset \mathcal{R}\{[\mathbf{S}_1(r)^T \ \cdots \ \mathbf{S}_P(r)^T]^T\}$;
2. $[\mathbf{S}_1(r+1)^T \ \cdots \ \mathbf{S}_P(r+1)^T]^T$ *is of full row rank,*

where $\mathcal{R}\{\cdot\}$ *denotes the row span. Then* $\mathbf{t}$ *must be a linear combination of* $\{\mathbf{s}_i\}$.

Proof:
See Appendix B. □

The above theorem concerns the identification of multiple sources. As the result indicates, the row span of the union of input matrices does not contain sufficient information to determine each individual input vector. However, it is adequate to identify the linear combination of $\{\mathbf{s}_i\}$. In other words, the span of the multiple input vectors $\mathcal{R}\{[\mathbf{s}_1 \cdots \mathbf{s}_P]^T\}$ can be determined from $\mathcal{R}\{[\mathbf{S}_1(r)^T \ \cdots \ \mathbf{S}_P(r)^T]^T\}$. Assuming that the channel matrix is of full column rank, $\mathcal{R}\{[\mathbf{s}_1 \cdots \mathbf{s}_P]^T\}$ can then be directly calculated from $\mathbf{X}$.

It can be shown that if $\mathbf{Y}$ is the null space of $\mathbf{V}$, it is related to $\mathbf{s}_1, \cdots, \mathbf{s}_P$ by

$$\mathbf{Y} = \mathbf{W}\left[\mathbf{s}_1 \cdots \mathbf{s}_P\right]^H,$$

where $\mathbf{W}$ is a $P \times P$ full rank matrix.

Without extra information, it does not seem to be possible to remove the ambiguity caused by $\mathbf{W}$. Fortunately, for most digital communication signals, the input symbols belong to a finite set of alphabets (e.g., BPSK, QPSK). It was proved in [11, 7] that blind symbol estimation can be achieved given sufficient data samples. The Iterative Least Squares with Projection (ILSP) algorithm introduced in [7] can be used to identify the input symbols from $\mathbf{Y}$.

5 Experimental Results

Field RF experiments were conducted to validate the proposed algorithm. In each experiment, an 8-element uniform linear antenna array was used to receive RF signals. The spatial oversampling rate is therefore 8. BPSK and QPSK signals of 50KHz baud rate were modulated to 900MHz and then transmitted from several transmitters located at the board side of the antenna array. In order to create a reliable flat fading scenarios, we used multiple transmitters to transmit delayed signals from different locations. The sampling rate is 50KHz, which was the same as the baud rate.

In the first experiment, a 4-ray multipath environment was generated by transmitting signals with relative delays at 0, 0.5, 1.2 and 1.6 baud periods from different locations. Figure 1 and 2 compare the signal constellations of one antenna outputs and the symbol estimates from the proposed algorithm using 50 sample vectors.

In the second experiment, we generated a multiple ($P = 2$) sources environment by letting two transmitters transmit $s_1(t)$ with relative delays at 0.1 and 0.7 baud periods, and another two transmitters transmit $s_2(t)$ with relative delays 0.2 and 0.8 baud periods, respectively. We applied the proposed estimation procedures for symbol estimation with 50 snapshots. In ten repeated experiments, the error probability of the symbol estimates was 0. In order to give a graphical illustration, we estimated the channels (a simple least-squares fitting) and then performed equalization the system outputs. The signal constellations are plotted in Figure 3.

Both experimental results were summarized in Table 1.

Appendix A Proof of Theorem 1

We use $\bar{\mathbf{a}}$ ($\underline{\mathbf{a}}$) to denote the vector $\mathbf{a}$ without the first (last) element. Define $\mathbf{s}_i$ and $\mathbf{t}_i$ the ith rows of two Hankel matrices $\mathbf{S}(r)$ and $\mathbf{T}(r)$, respectively.

Given the row span of $\mathbf{S}(r)$, if there is another Hankel matrix $\mathbf{T}(r)$ such that $\mathcal{R}\{\mathbf{T}(r)\} \subset \mathcal{R}\{\mathbf{S}(r)\}$, then

$$\mathbf{T}(r) = \mathbf{W}\mathbf{S}(r), \tag{A.1}$$

where $\mathbf{W}$ is a $r \times r$ non-zero matrix.

For any consecutive rows in the above equation ($i = 2, \cdots, r$), we have

$$\begin{aligned} \mathbf{t}_{i-1} &= \sum_{k=1}^{r} \mathbf{W}(i-1,k)\mathbf{s}_k, \qquad (A.2) \\ \mathbf{t}_i &= \sum_{k=1}^{r} \mathbf{W}(i,k)\mathbf{s}_k, \end{aligned}$$

where $\mathbf{W}(i,j)$ is the element of $\mathbf{W}$ at the ith row and jth column. Consequently,

$$\begin{aligned} \underline{\mathbf{t}}_{i-1} &= \sum_{k=1}^{r} \mathbf{W}(i-1,k)\underline{\mathbf{s}}_k, \qquad (A.3) \\ \bar{\mathbf{t}}_i &= \sum_{k=1}^{r} \mathbf{W}(i,k)\bar{\mathbf{s}}_k. \end{aligned}$$

Using the fact that $\bar{\mathbf{t}}_i = \underline{\mathbf{t}}_{i-1}$ and $\bar{\mathbf{s}}_i = \underline{\mathbf{s}}_{i-1}$, we obtain

$$\mathbf{W}(i,1)\bar{\mathbf{s}}_1+\sum_{k=2}^{r}(\mathbf{W}(i,k)-\mathbf{W}(i-1,k-1))\bar{\mathbf{s}}_k+\mathbf{W}(i-1,r)\bar{\mathbf{s}}_r = \mathbf{0}. \quad (A.4)$$

Under the assumption that $\mathbf{s}$ contains more than r modes, the rows of $\mathbf{S}(r+1)$ must be linear independent [3]. Therefore, all the coefficients in the above equations are zeros,

$$\begin{aligned} \mathbf{W}(i,1) &= 0; \\ \mathbf{W}(i,k) &= \mathbf{W}(i-1,k-1), \quad k = 2,\cdots,r-1; \\ \mathbf{W}(i-1,r) &= 0. \end{aligned} \quad (A.5)$$

Consider all the possible i values, $\mathbf{W} = \alpha\mathbf{I}$. In other words, $\mathbf{t} = \alpha\mathbf{s}$. □

Appendix B Proof of Theorem 2

Similar to (A.1), we now have

$$\mathbf{T}(r) = \sum_{i=1}^{P} \mathbf{W}_i \mathbf{S}_i(r).$$

Under the assumption, *i.e.*, $[\mathbf{S}_1(r+1)^T, \cdots, \mathbf{S}_P(r+1)^T]^T$ is of full row rank, the same argument in the proof of Theorem 1 follows, which leads to

$$\mathbf{W}_i = \alpha_i \mathbf{I}, \quad i = 1, \cdots, P.$$

Thus, $\mathbf{T}(r) = \sum_{i=1}^{P} \alpha_i \mathbf{S}_i(r)$. Comparing each element of the two matrices on both sides, we obtain $\mathbf{t} = \sum_{i=1}^{P} \alpha_i \mathbf{s}_i$.

References

[1] L. A. Baccala and S. Roy. A new blind time-domain channel identification method based on cyclostationarity. *To appear in IEEE Signal Processing Letters*, 1994.

[2] Z. Ding. Blind identification and equalization using spectral correlation measurements: Part I – frequency domain analysis. In W. A. Gardner, editor, *Cyclostationarity in communications and signal processing*. IEEE Press, 1993.

[3] H. Liu, G. Xu, and L. Tong. "A Deterministic Approach to Blind Equalization". In *Proc. Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 1993.

[4] E. Moulines, P. Duhamel, J. Cardoso, and S. Mayrargue. Subspace methods for the blind identification of multichannel FIR filters. In *Proc. IEEE ICASSP'94*, pages IV573–IV576, April 1994.

[5] S. V. Schell, D. L. Smith, and S. Roy. Blind channel identification using subchannel response matching. In *Proc. 1994 Conf. Information Sciences and Systems*, Princeton, NJ, March 1990.

[6] D. T. M. Slock. Blind fractionally-spaced equalization, perfect-reconstruction filter banks and multichannel linear prediction. In *Proc. IEEE ICASSP'94*, pages IV585–IV588, April 1994.

[7] S. Talwar, M. Viberg, and A. Paulraj. Blind estimation of multiple co-channel digital signals using an antenna array. *IEEE Signal Processing Letters*, 1(2):29–31, February 1994.

[8] L. Tong, G. Xu, and T. Kailath. "A new approach to blind identification and equalization of multipath channel". In *Proc. of the 25^{th} Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November 1991.

[9] J. K. Tugnait. On blind identifiability of multipath channels using fractional sampling and second-order cyclostationary statistics. In *Proc. Global Telecom. Conf.*, pages 2000–2005, 1993.

[10] G. Xu, H. Liu, L. Tong, and T. Kailath. Deterministic blind identification of multichannel FIR systems. *submitted to IEEE Trans. Signal Processing*, 1993.

[11] D. Yellin and B. Porat. Blind identification of fir systems excited by discrete-alphabet inputs. *IEEE Trans. Signal Processing*, SP-41(3):1331–1339, 1993.

	Modu.	# of sources	# of paths
Exp. 1	QPSK	1	4
Exp. 2	BPSK	2	2; 2

	delays [T]	RMSE (before)	RMSE (after)
Exp.1	0, 0.5, 1.2, 1.6	18.57	0.1129
Exp. 2	0.1, 0.7; 0.2, 0.8	N/A	0.0365; 0.0747

Table 1: Experimental Results

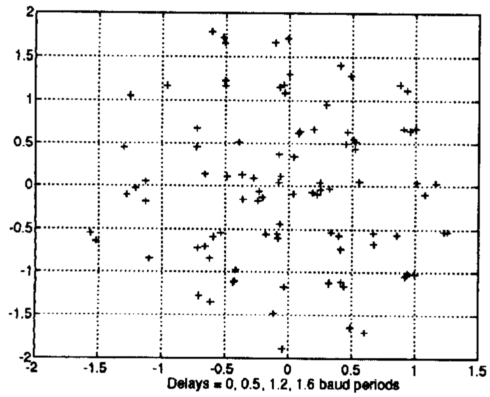

Figure 1: Signal Constellation of Antenna Outputs

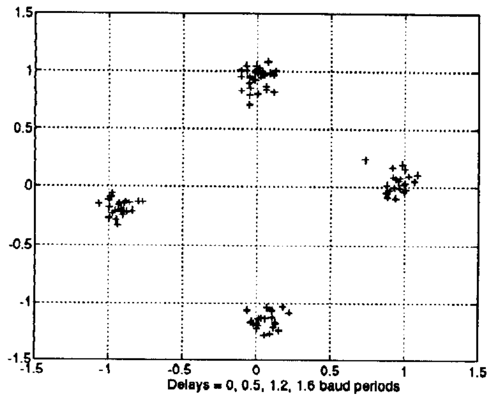

Figure 2: Signal Constellation after Processing

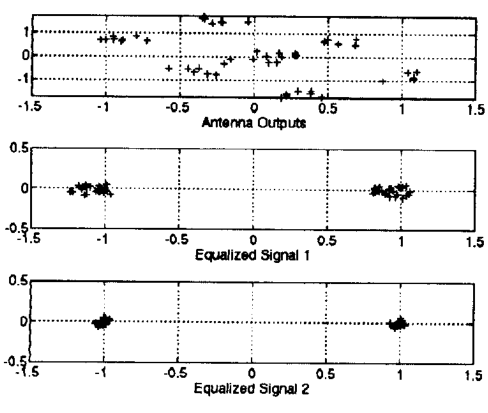

Figure 3: Signal Constellation: Multiple Sources