

BLIND IDENTIFICATION OF FIR CHANNELS CARRYING MULTIPLE FINITE ALPHABET SIGNALS

Alle-Jan van der Veen^{*†}, Shilpa Talwar[†] and Arogyaswami Paulraj^{*}

^{*}Information Systems Laboratory/ [†]Scientific Computing Program,
Stanford University, Stanford, CA 94305

Abstract

The finite alphabet property of digital communication signals, along with oversampling techniques, enables the blind identification and equalization of an unknown FIR channel carrying a superposition of such signals, provided they have the same (known) period. Applied to multi-user wireless communications, the same framework allows the blind separation of multiple finite alphabet signals received at an arbitrary antenna array through an unknown multipath propagation environment with finite delay spread. An algorithm is proposed and tested on simulated data.

1. INTRODUCTION

In the context of blind identification of channels carrying digital communication signals, a number of algorithms have been proposed to estimate

(A) a single FIR channel carrying one signal (FIR-SISO).

In one class of algorithms, initiated by Tong, Xu and Kailath [1], the signal is recovered by oversampling the channel output (viz. a.o. [2–6]). There are many other blind equalization methods, for example based on high-order statistics.

In the context of array signal processing, another scenario which admits blind identification is the case where

(B) M antennas receive a superposition of $d \leq M$ synchronized finite-alphabet (FA) input signals via memoryless channels.

One algorithm to recover the signals was recently proposed by Talwar, Viberg and Paulraj [7, 8]. In the present paper, we combine the above two scenarios and derive an algorithm to

(C) blindly identify multiple FIR channels carrying a superposition of unsynchronized digital FA input signals that have the same symbol rate and alphabet (FIR-MISO or FIR-MIMO case). See figure 1.

In a deterministic setting, the FIR-MIXO case has not yet received much attention. Although a few adaptive antenna combining/equalizing algorithms have been proposed [9, 10], these require long data runs for convergence and are also not satisfying from a theoretical perspective.

Our algorithm for (C) consists of two steps. The first step is an extension from scalars to vectors of an algorithm for scenario (A) to handle more than one signal. At this point, the ISI caused by the channel is removed and the input signals are synchronized. However, the symbol sequences can be determined only up to a fixed linear combination of them. This is precisely scenario (B).

This research was partly supported by ARPA, contract no. F49620-91-C-0086, monitored by the AFOSR. A.J. van der Veen is presently with Delft Univ. of Techn., Dept. EE, 2628 CD Delft, The Netherlands.

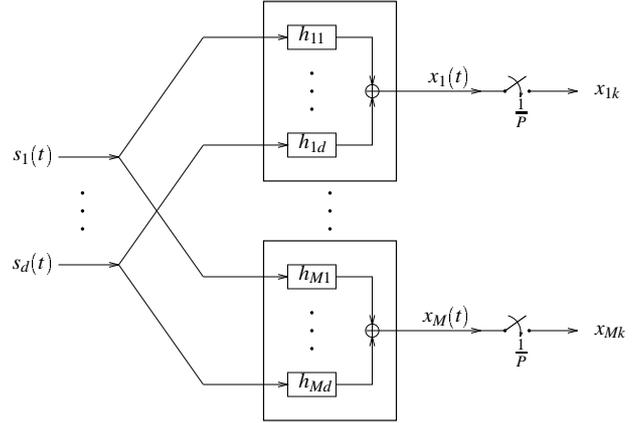


Figure 1. Channel model.

A multiple-sensor version (the FIR-MIMO case) is covered by the same algorithm. The additional sensors improve the conditioning of the problem, in particular in the case of bandlimited signals and high noise levels. In the context of array signal processing, the algorithm can be used to separate a number of incoming digital signals, arriving from different or possibly the same directions, and distorted by multipath with finite delay spread. In theory, *i.e.*, under low noise conditions, the only fundamental restriction is that the number of antennas times the oversampling factor should be larger than the number of independent signals. The required sampling period is typically in the order of 50 symbol periods.

2. DATA MODEL

We describe a digital signal $s(t)$ as a sequence of dirac pulses, $s(t) = \sum_{k=-\infty}^{\infty} s_k \delta(t - k)$. For convenience, the symbol rate is normalized to $T = 1$. An array of M sensors, with outputs $x_1(t), \dots, x_M(t)$, receives d digital signals $s_1(t), \dots, s_d(t)$ through independent channels $h_{ij}(t)$. Each impulse response $h_{ij}(t)$ is a convolution of the shaping filter of the i -th signal and the actual channel from the i -th input to $x_j(t)$, including propagation delays and delays that allow to model unsynchronous signals as synchronous. The data model is written compactly as the convolution $\mathbf{x}(t) = H(t) * \mathbf{s}(t)$, where

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_M(t) \end{bmatrix}, H(t) = \begin{bmatrix} h_{11}(t) & \cdots & h_{1d}(t) \\ \vdots & & \vdots \\ h_{M1}(t) & \cdots & h_{Md}(t) \end{bmatrix}, \mathbf{s}(t) = \begin{bmatrix} s_1(t) \\ \vdots \\ s_d(t) \end{bmatrix}$$

If we assume that all $h_{ij}(t)$ are FIR filters of length at most $L \in \mathbb{N}$:

$$h_{ij}(t) = 0, \quad t \notin [0, L), \quad (1)$$

then at most L consecutive symbols of each signal play a role in $\mathbf{x}(t)$ at any given moment: for $t = n + \tau$, where $n \in \mathbb{Z}$ and $0 \leq \tau < L$,

$$x_i(n + \tau) = \sum_{k=0}^{L-1} h_{i1}(k + \tau) s_{1,n-k} + \dots + \sum_{k=0}^{L-1} h_{id}(k + \tau) s_{d,n-k}.$$

Further suppose that each $x_i(t)$ is sampled at a rate $P \in \mathbb{N}$, where P is the oversampling factor. If we start sampling at time $t = 0^*$ and collect samples during N symbol periods, then we can construct a data matrix X as

$$X = \begin{bmatrix} \mathbf{x}_0 & \dots & \mathbf{x}_{N-1} \\ \mathbf{x}(0) & \mathbf{x}(1) & \dots & \mathbf{x}(N-1) \\ \mathbf{x}(\frac{1}{P}) & \mathbf{x}(1 + \frac{1}{P}) & \dots & \dots \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}(\frac{P-1}{P}) & \dots & \dots & \mathbf{x}(N-1 + \frac{P-1}{P}) \end{bmatrix}.$$

The k -th column \mathbf{x}_k of X contains the MP samples taken in the k -th symbol period. With the model of $x_i(t)$, it follows that X has a factorization

$$X = HS_T$$

where

$$H := \begin{bmatrix} H(0) & H(1) & \dots & H(L-1) \\ H(\frac{1}{P}) & \dots & \dots & \dots \\ \vdots & \vdots & \ddots & \vdots \\ H(\frac{P-1}{P}) & \dots & \dots & H(L-1 + \frac{P-1}{P}) \end{bmatrix}$$

$H : MP \times dL$,

$$S_T := \begin{bmatrix} \mathbf{s}_0 & \dots & \mathbf{s}_{N-2} & \mathbf{s}_{N-1} \\ \vdots & \vdots & \vdots & \mathbf{s}_{N-2} \\ \mathbf{s}_{-L+2} & \mathbf{s}_{-L+3} & \dots & \dots \\ \mathbf{s}_{-L+1} & \mathbf{s}_{-L+2} & \dots & \mathbf{s}_{N-L} \end{bmatrix}$$

$S_T : dL \times N$, block-Toeplitz.

The blind identification problem is to estimate H and S_T from X . Note that for such a factorization to be unique, it is necessary that H and S_T have full column rank and row rank, respectively, which implies a.o. $MP \geq dL$. If this condition does not hold, we can extend X to a block-Hankel matrix, by left-shifting and stacking m times,

$$\mathcal{X} = \begin{bmatrix} \mathbf{x}_0 & \mathbf{x}_1 & \dots & \mathbf{x}_{N-m} \\ \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_{N-2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_{m-1} & \dots & \mathbf{x}_{N-2} & \mathbf{x}_{N-1} \end{bmatrix}$$

$\mathcal{X} : mMP \times (N - m + 1)$.

The augmented data matrix \mathcal{X} has a factorization

$$\mathcal{X} = \mathcal{H}\mathcal{S}$$

$$= \begin{bmatrix} \mathbf{0} & H \\ H & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{s}_{m-1} & \dots & \mathbf{s}_{N-2} & \mathbf{s}_{N-1} \\ \vdots & \vdots & \vdots & \mathbf{s}_{N-2} \\ \mathbf{s}_{-L+2} & \mathbf{s}_{-L+3} & \dots & \dots \\ \mathbf{s}_{-L+1} & \mathbf{s}_{-L+2} & \dots & \mathbf{s}_{N-L-m+1} \end{bmatrix}$$

$\mathcal{H} : mMP \times d(L + m - 1)$: block-Hankel,
 $\mathcal{S} : d(L + m - 1) \times (N - m + 1)$: block-Toeplitz.

*This is for notational convenience and without loss of generality.

Now, necessary conditions for \mathcal{X} to have a minimal-rank factorization $\mathcal{X} = \mathcal{H}\mathcal{S}$ are that \mathcal{H} is a ‘tall’ matrix and \mathcal{S} is a ‘wide’ matrix, which for $L > 1$ leads to

$$\begin{aligned} MP &> d \\ m &\geq \frac{dL - d}{MP - d} \\ N &> dL + (d + 1)(m - 1). \end{aligned} \quad (2)$$

Only $MP > d$ is a fundamental restriction.

3. BLIND IDENTIFICATION

Suppose that the conditions (2) are satisfied. Then

$$\begin{aligned} \mathcal{H} \text{ full column rank} &\Rightarrow \text{row}(\mathcal{X}) = \text{row}(\mathcal{S}) \\ \mathcal{S} \text{ full row rank} &\Rightarrow \text{col}(\mathcal{X}) = \text{col}(\mathcal{H}) \end{aligned}$$

To factor \mathcal{X} into $\mathcal{X} = \mathcal{H}\mathcal{S}$, the strategy is to find either \mathcal{S} : a block-Toeplitz matrix with a specified row span, or \mathcal{H} : a block-Hankel matrix with a specified column span. In the scalar case ($d = 1$ signal), a number of algorithms have been proposed for doing this [1–6]. It is straightforward to extend these algorithms to the vector case ($d > 1$). However, for $d > 1$ subspace information alone leads to an ambiguity: $\mathcal{X} = (\mathcal{H}D^{-1})(D\mathcal{S})$ is a factorization with the same subspaces, for $D = \text{diag}[A, \dots, A]$ and A any invertible $d \times d$ matrix. This ambiguity is resolved in a second step, by taking advantage of the finite-alphabet property of the signals.

3.1. Direct estimation of \mathcal{S}

A standard procedure to find \mathcal{S} as a block-Toeplitz matrix with $\text{row}(\mathcal{S}) = \text{row}(\mathcal{X})$ is to rewrite these conditions as

$$\begin{aligned} [\mathbf{s}_{-L+1} \quad \mathbf{s}_{-L+2} \quad \dots \quad \mathbf{s}_{N-L-m+1}] &\in \text{row}(\mathcal{X}) \\ \vdots & \\ [\mathbf{s}_{m-1} \quad \mathbf{s}_m \quad \dots \quad \mathbf{s}_{N-1}] &\in \text{row}(\mathcal{X}). \end{aligned}$$

Hence, $S := [\mathbf{s}_{-L+1} \quad \mathbf{s}_{-L+2} \quad \dots \quad \mathbf{s}_{N-1}]$ is in the *intersection* of the rowspan of \mathcal{X} and shifts of this row span (suitably embedded with zeros). Alternatively, we can say that S is orthogonal to the *union* of the complement of these row spans. The latter space might be easier to construct, except perhaps for large N . Thus let G be a matrix whose columns constitute a basis for $\ker(\mathcal{X})$. If \mathcal{H} has full column rank, then G has dimensions $(N - m + 1) \times (N - m + 1 - d(L + m - 1)) =: m_G \times N_G$. Moreover, $\mathcal{X}G = 0 \Rightarrow SG = 0$. Using the fact that \mathcal{S} is block-Toeplitz, we obtain

$$SG = 0 \Leftrightarrow SG_{T(\ell)} = 0,$$

$$G_{T(\ell)} := \begin{bmatrix} G & & & & \\ & G & & & \\ & & \ddots & & \\ & & & G & \\ \mathbf{0} & & & & \end{bmatrix}$$

$G_{T(\ell)} : (N + \ell - 1) \times N_G(\ell + m - 1)$.

The number of block-columns of $G_{T(\ell)}$ is equal to $\ell + m - 1$, where ℓ is a parameter chosen equal to the channel length L (or maybe smaller, viz. section 3.5).

If $G_{T(\ell)}$ is a wide matrix (this gives additional conditions on m and N), then $\ker(G_{T(\ell)}^*)$ determines S , but only up to a left invertible $d \times d$ matrix A , because $Y = AS$ also satisfies $YG_{T(\ell)} = 0$.

Given $G_{T(L)}$, we take Y to be a matrix whose rows form a basis for $\ker(G_{T(L)}^*)$. To identify S , we have to find the factorization $Y = AS$, which, in the case of finite alphabet signals, can be done using the ILSP algorithm.

3.2. The ILSP algorithm

For a given Y , the ILSP algorithm [7, 8] solves the factorization ($Y = AS : A, S$ full rank, $[S]_{ij} \in \mathcal{FA}$), where \mathcal{FA} is a pre-specified finite alphabet. In its simplest formulation, the algorithm consists of alternating projections: starting, e.g., with $S^{(0)} = Y$,

- Project $S^{(k)}$ onto $\text{row}(Y)$: $S^{(k)'} := S^{(k)}Y^\dagger Y$,
- Project each $[S^{(k)'}]_{ij}$ onto the closest member of the alphabet, resulting in $S^{(k+1)}$.

The iteration generally converges very rapidly. Note that if we took Y to be an *orthonormal* basis, then $Y^\dagger = Y^*$.

3.3. Computation of \mathcal{H} first

Instead of estimating \mathcal{S} directly, we can also first estimate \mathcal{H} and invert the resulting channel to estimate \mathcal{S} . This is potentially more interesting, since dimensions do not grow with N .

The approach we take here is basically that of [5]. Let G' be a basis of the left kernel of \mathcal{X} . Assuming \mathcal{S} to be of full rank, we have $G'\mathcal{X} = 0 \Rightarrow G'\mathcal{H} = 0$. Write

$$\begin{aligned} H &:= [H_0 \ \cdots \ H_{L-1}], \quad H_i : P \times d \\ G' &:= [G'_1 \ \cdots \ G'_m], \quad G'_i : m_{G'} \times P \\ \mathcal{H} &= \begin{bmatrix} \mathbf{0} & & H_0 & H_1 & \cdots & H_{L-1} \\ & \ddots & & & & \\ H_0 & H_1 & \cdots & H_{L-1} & & \mathbf{0} \end{bmatrix} \end{aligned}$$

Then $G'\mathcal{H} = 0 \Leftrightarrow$

$$\begin{bmatrix} G'_m & & \mathbf{0} \\ \vdots & \ddots & \\ G'_1 & & G'_m \\ \mathbf{0} & & G'_1 \end{bmatrix} \begin{bmatrix} H_0 \\ \vdots \\ H_{L-1} \end{bmatrix} = 0.$$

This specifies H , up to a right block-diagonal factor $\text{diag}[A, \dots, A]$. S is found as $S = \mathcal{H}^\dagger \mathcal{X}$, and is block-Toeplitz in the no-noise case. With noise, we can average along the diagonals of \mathcal{S} to obtain a Toeplitz structure. At this point, the ILSP algorithm is employed to remove the ambiguity that A represents.

For the estimation of \mathcal{H} , it is only required that \mathcal{S} be of full row rank, which is a mild condition. In particular, it is not necessary that all channels have equal length. In general, estimating \mathcal{H} is computationally easier (for large N), but our experience with simulations is that estimating \mathcal{S} directly might be more accurate.

3.4. Detection of d and L

If \mathcal{H} and \mathcal{S} have full column rank and row rank, respectively, then the rank of \mathcal{X} is $d_{\mathcal{X}} := d(L + m - 1)$. In principle, the number of signals d can be estimated by increasing the blocking factor m of \mathcal{X} by one, and looking at the increase in rank of \mathcal{X} . This property provides a useful detection mechanism even if the noise level is quite high since it is independent of the actual (observable) channel length \hat{L} . Furthermore, it still holds if all channels do not have equal lengths (see section 3.5 below). If they do, then L can be estimated from the estimated rank of \mathcal{X} , $\hat{d}_{\mathcal{X}}$, and the estimated number of signals, \hat{d} , by $\hat{L} = \hat{d}_{\mathcal{X}} / \hat{d} - m + 1$.

3.5. Unequal channel lengths

If the channels do not have equal length, but lengths L_{ij} , say, then \mathcal{H} is not full rank and a modification of the algorithm for estimating \mathcal{S} is necessary. Define $L_j = \max_i L_{ij}$, the maximum number of symbols of signal j that play a role in a single sample vector $\mathbf{x}(t)$. Furthermore, define the overall channel length as $L = \max_j L_j$, which is the same as before in (1). Then H has only $\sum_{j=1}^d L_j$ columns that are not identical to zero. The remaining $dL - \sum_j L_j$ zero columns show up in the right block columns of H , so that, generically,[†] $\text{rank}(\mathcal{H})$ is also reduced to $(\sum_j L_j) + d(m - 1)$. Hence, the rank of \mathcal{X} is $\text{rank}(\mathcal{X}) = (\sum_j L_j) + d(m - 1)$. (Note that it is still possible to detect d by increasing m by one.) Since certain columns of \mathcal{H} are zero, the corresponding rows of \mathcal{S} are not in $\text{row}(\mathcal{X})$, and the corresponding signals do not satisfy $sG_{T(L)} = 0$. Hence $\dim \ker(G_{T(L)}^*) = \#(L_j \equiv L)$ is equal to the number of signals for which $L_j = L$, and only these signals are obtained after processing the basis of the kernel by ILSP. We can, however, underestimate L by \hat{L} , say, and thus take less row span intersections, or less shifts in $G_{T(\hat{L})}$. Then

$$\hat{d} := \dim \ker(G_{T(\hat{L})}^*) = \#(L_j \geq \hat{L}) + \#(L_j \geq \hat{L} + 1) + \cdots + \#(L_j \geq L).$$

In this case, signals for which $L_j \geq \hat{L}$ are represented in $\ker(G_{T(\hat{L})})$, as well as their shifts over up to $(L_j - \hat{L})$ positions. If we take $\hat{L} = \min L_j$, then all signals are represented at least once in the kernel.

As before, the \hat{d} basis vectors in the kernel of $G_{T(\hat{L})}$ are processed by ILSP to remove the ambiguity in the choice of the basis. At this point, it is straightforward to detect whether a signal is a shifted version of another signal in the collection, in which case one of them has to be discarded. For $\hat{L} = \min L_j$, the signals that remain are the d independent signals.

3.6. Remarks

The above approach of underestimating L and overestimating d appears to make the overall algorithm more robust in the presence of noise as well, also in the case that all signals do have the same channel length. The reason is that ILSP gets a larger responsibility in separating the signals and their echos, which is favorable because the finite alphabet property is quite powerful.

Singular value decompositions are used to estimate subspaces in the presence of noise. Fast subspace techniques which estimate only a few singular vectors are clearly a method of choice in the estimation of $\dim \ker(G_{T(\hat{L})}^*)$.

4. SIMULATION RESULTS

To demonstrate the viability of the identification scheme by an example, we consider the following multiray scenario. In the simulation, $d = 2$ BPSK signals are broadcast from certain locations, each modulated by a raised cosine waveform $W(t) = \sin(t)/t \cdot \cos(\beta t)/(1 - 2\beta t)$, truncated at a length of 6 baud periods, and with modulation parameter $\beta = 0.35$. The signals are received by $M = 2$ identical omnidirectional antennas, spaced by half a wavelength. The simulated channel consists of four paths per signal, where each path is specified by an angle-of-arrival α , delay τ , and complex

[†]Here, ‘generically’ means provided the overall channel is otherwise identifiable, i.e., unless the channels have ‘common zeros’ in the sense of [3]. Precise identifiability conditions are beyond the scope of the paper.

Table 1. Bit error rates, standard deviations, and number of times a signal is not detected (over 50 runs).

$m = 3$	\hat{d}_X	$\hat{L} = 2$				$\hat{L} = 3$				$\hat{L} = 4$									
		BERs (%)		STDs		BERs (%)		STDs		BERs (%)		STDs		ND					
11	2	8.86	15.27	0.53	0.49	2	13	8.58	24.73	0.52	0.56	2	17	13.94	20.06	0.57	0.68	7	9
	4	4.37	5.22	0.39	0.38	1	2	3.96	6.73	0.42	0.43	0	2	8.62	12.09	0.58	0.63	1	1
	6	2.47	1.25	0.30	0.26	0	0	3.38	5.19	0.39	0.42	0	0	7.40	9.51	0.56	0.58	0	0
12	2	8.31	13.51	0.52	0.48	2	9	8.27	25.69	0.51	0.58	3	19	12.04	14.66	0.55	0.63	6	3
	4	3.41	5.24	0.39	0.38	0	1	3.31	4.04	0.38	0.40	1	1	7.00	8.45	0.52	0.57	1	0
	6	2.24	2.20	0.29	0.27	0	0	3.08	2.92	0.33	0.35	0	0	4.58	6.98	0.46	0.52	1	0
13	2	8.55	12.65	0.53	0.48	2	9	8.42	25.65	0.51	0.58	2	19	9.21	13.94	0.56	0.62	2	3
	4	3.75	3.37	0.40	0.37	1	0	3.27	3.46	0.38	0.42	1	0	5.89	7.96	0.49	0.54	0	0
	6	2.27	0.86	0.30	0.25	0	0	2.35	2.73	0.32	0.34	0	0	4.75	6.53	0.46	0.49	0	0
14	2	8.31	13.94	0.53	0.47	2	11	8.88	26.27	0.52	0.58	3	20	12.75	13.74	0.56	0.61	6	2
	4	3.20	3.94	0.40	0.37	1	1	2.31	2.62	0.38	0.40	0	0	6.55	7.96	0.53	0.53	1	0
	6	2.16	1.29	0.30	0.24	0	0	2.50	2.23	0.34	0.33	0	0	4.23	5.85	0.45	0.46	0	0

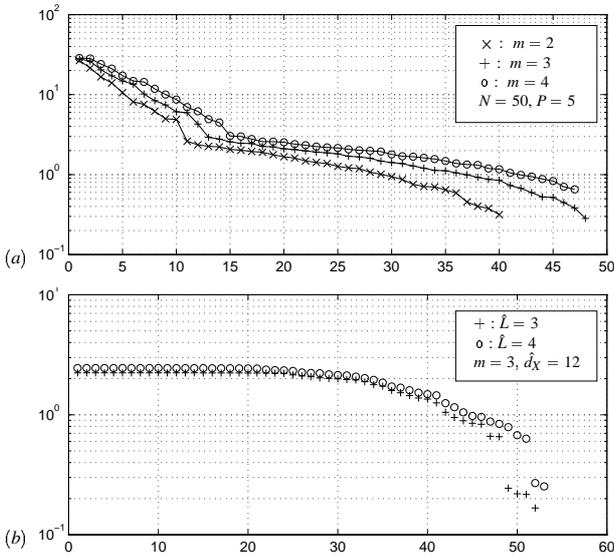


Figure 2. (a) Singular values of \mathcal{X} , for different blocking factors m . (b) Singular values of $G_{T(\hat{L})}$.

damping factor p , as by the following table:

Signal 1:	α :	-10°	-2°	-120°	160°
	τ :	0	0.3	1.2	2.1
	$ p $:	1	0.8	0.4	0.4
Signal 2:	α :	10°	15°	-40°	150°
	τ :	0.5	0.9	1.5	2.8
	$ p $:	1	0.9	0.5	0.3

The resulting channel length is $L = 9$. The complex phase of each damping factor p was selected randomly from $[0, 2\pi]$ with a uniform distribution. We took $N = 50$ sampling intervals, $P = 5$ times oversampling, and $\text{SNR} = 10$ dB per signal.

The singular values of one instance of the data matrix \mathcal{X} are plotted in figure 2(a), for blocking factors $m = 2, 3, 4$. The estimated numerical rank of \mathcal{X} is about $\hat{d}_X = 10, 12, 14$, respectively, so that, with $d_X = d(L + m - 1)$, the number of signals is detected as $\hat{d} = 2$, and the channel length is detected as $\hat{L} = 4$, rather than 9. Figure 2(b) shows the singular values of $G_{T(\hat{L})}$, for $m = 3$, $\hat{d}_X = 12$, and $\hat{L} = 3, 4$. For $\hat{L} = 4$, the number of small singular values is in-

deed equal to $\hat{d} = 2$, for $\hat{L} = 3$, we also obtain shifted copies of the two signals. Table 1 gives the bit-error rates and symbol standard deviations (before classification as +1 or -1), for various choices of \hat{L} , \hat{d}_X and \hat{d} , and averaged over 50 independent runs. We also list how often a signal was *not* in the kernel (ND). It is seen that choosing \hat{L} smaller than estimated, in combination with overestimating \hat{d} so that more singular vectors are processed by the ILSP algorithm, leads to an important decrease in bit errors and parameter standard deviations. The choice of \hat{d}_X appears to be not critical.

5. REFERENCES

- [1] L. Tong, G. Xu, and T. Kailath, "A new approach to blind identification and equalization of multipath channels," in *25-th Asilomar Conf. on Signals, Systems and Computers*, pp. 856–860 vol.2, 1991.
- [2] H. Liu, G. Xu, and L. Tong, "A deterministic approach to blind equalization," in *27-th Asilomar Conf. on Signals, Systems and Computers*, 1993.
- [3] L. Tong, G. Xu, and T. Kailath, "Fast blind equalization via antenna arrays," *Proc. IEEE ICASSP*, pp. IV:272–274, 1993.
- [4] D. Slock, "Blind fractionally-spaced equalization, perfect-reconstruction filter banks and multichannel linear prediction," in *Proc. IEEE ICASSP*, pp. IV:585–588, 1994.
- [5] E. Moulines, P. Duhamel, J. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *Proc. IEEE ICASSP*, pp. IV:573–576, 1994.
- [6] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Tr. Inform. Th.*, vol. 40, pp. 340–349, 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*, vol. 1, pp. 29–31, Feb. 1994.
- [8] S. Talwar and A. Paulraj, "Performance analysis of blind digital signal copy algorithms," in *Proc. IEEE MILCOM*, pp. 123–127 vol.1, 1994.
- [9] J.R. Treichler and M.G. Larimore, "New processing techniques based on constant modulus adaptive algorithm," *IEEE Tr. Acoust., Speech, Signal Proc.*, vol. 33, pp. 420–431, 1985.
- [10] R.P. Gooch and B.J. Sublett, "Joint spatial and temporal equalization in a decision-directed adaptive antenna system," in *22-nd Asilomar Conf.*, pp. 255–259 vol.1, 1988.