A Fast Blind SIMO Channel Identification Algorithm for Sparse Sources

David Luengo, *Student Member, IEEE*, Ignacio Santamaría, *Member, IEEE*, Jesús Ibáñez, Luis Vielva, and Carlos Pantaleón, *Member, IEEE*

***Abstract*—We address the blind identification of single-input-multiple output (SIMO) finite impulse response systems when the input signal is sparse. The problem is equivalent to underdetermined blind source separation (BSS), but with temporal correlation among the sources. Exploiting the sparse character of the input signal, the algorithm solves three different problems: first, to estimate the directions of the columns of the channel matrix; second, to estimate the L_2-norm of the columns; and finally, to find the correct ordering of the columns of the mixing matrix. The last step is not required for the blind source separation (BSS) problem, since any permutation of the columns is admissible for BSS. The performance and computational cost of the algorithm in a noiseless situation is compared against subspace-based techniques.**

***Index Terms*—Blind channel identification, blind source separation (BSS), sparse deconvolution.**

I. Introduction

BLIND channel identification is an important and widely studied problem that appears in many signal processing applications: equalization, seismic data deconvolution, speech coding, image deblurring, etc. When the channel impulse response has finite support and the received signal is oversampled, the problem can be formulated as the blind identification of a single-input multiple-output (SIMO) finite-impulse response (FIR) system. In this situation, it is known that, as long as the FIR channels have no common zeros and the channel is fully excited, the SIMO system can be identified using only second-order statistics of the output [1]. Several extensions of this idea using subspace-based methods [2] or linear prediction techniques [3] have proven useful to solve this problem. An iterative algorithm based on second-order statistics is presented in [4] where the order of the system is also estimated for channels with rational transfer functions. The main drawback of all these algorithms, however, is their high computational cost. Therefore, there is still a need for low-cost, fast algorithms that can deal with the large datasets typically available in some applications such as seismic deconvolution or nondestructive evaluation.

Since, typically, the number of columns of the SIMO channel matrix is larger than the number of measurements, blind SIMO identification can be viewed as a problem of underdetermined blind source separation (i.e., more sources than sensors), but with temporal correlation among the sources. On the other hand, recent advances in BSS have shown that, when the sources are sparse, it is possible to solve the underdetermined case using simple algorithms [5]–[7]. The general idea of these techniques is that, when the sources are sparse, the measurements tend to cluster along the directions imposed by the columns of the mixing matrix, which can then be easily identified.

Using similar ideas, in this letter we propose a fast blind SIMO channel identification algorithm for sparse sources with application to seismic deconvolution, nondestructive evaluation, or blurred star image deconvolution. Exploiting the sparsity of the input signal, a blind technique must solve three different problems. First, it is necessary to estimate the directions of the columns of the channel matrix; once the directions have been identified, the norm of the columns must be estimated. These two steps solve the blind source separation (BSS) problem, since any permutation of the columns of the mixing matrix is admissible. However, in a deconvolution problem the ordering of the columns of the channel matrix is also required. In this letter we describe a simple algorithm to solve these three steps.

II. Problem Statement

Assuming that the received signal is oversampled by a factor m and that the maximum length of each of the FIR channels is $q > m$; then, in a noiseless situation, the problem can be formulated as

$$\mathbf{X} = \mathbf{H}\mathbf{S} \tag{1}$$

where $\mathbf{X} = [\mathbf{x}(1) \cdots \mathbf{x}(N)]$ is an $m \times N$ matrix, formed by stacking N successive observations $\mathbf{x}(n) = [x_1(n) \cdots x_m(n)]^T$; $\mathbf{S} = [\mathbf{s}(1) \cdots \mathbf{s}(N)]$ is a $q \times N$ matrix of input signals, with $\mathbf{s}(n) = [s(n) \cdots s(n-q+1)]^T$; and $\mathbf{H} = [\mathbf{h}(0) \cdots \mathbf{h}(q-1)]$ is an $m \times q$ channel or mixing matrix, with $\mathbf{h}(k) = [h_1(k) \cdots h_m(k)]^T$.

Our problem consists of estimating the channel matrix $\mathbf{H}$ using only the observations and some statistical knowledge of the input signal. Specifically, in this letter, we assume that the input signal $s(n)$ is a sparse spike train, modeled as a statistically independent zero-mean Bernoulli–Gaussian (BG) sequence, which is commonly used in nondestructive evaluation and seismic deconvolution [8]. The BG samples are generated according to

$$f_S(s) = p\delta(s) + (1-p)\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-s^2}{2\sigma^2}\right) \tag{2}$$

Manuscript received April 19, 2002; revised September 3, 2003. This work was supported in part by MCYT (Ministerio de Ciencia y Tecnología) under Grant TIC2001-0751-C04-03. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Petr Tichavsky.

The authors are with the Department of Communications Engineering (DICOM), Universidad de Cantabria, 39005 Santander, Spain (e-mail: nacho@gtas.dicom.unican.es).

Digital Object Identifier 10.1109/LSP.2003.810014

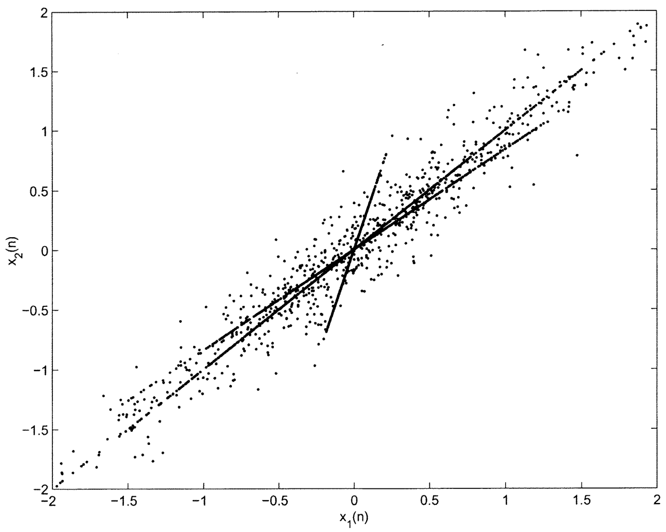

Fig. 1. Scatter plot of the observations for $m = 2, q = 3, N = 10\,000$ input samples and a sparsity factor $p = 0.75$.

where p is the sparsity factor, and σ^2 the variance of the nonzero samples. Moreover, since in blind system identification there is always a global scale factor ambiguity, the variance can be arbitrarily chosen as $\sigma^2 = 1$.

The nth output vector in (1) can be expressed as a linear combination of the columns of the channel matrix

$$\mathbf{x}(n) = \sum_{k=0}^{q-1} \mathbf{h}(k) s(n-k). \tag{3}$$

Thus, if only one sample of the BG sequence is nonzero for $n, \ldots, n-q+1$, the measurement at time n will be collinear with the corresponding column of the channel matrix. This fact has been exploited in several ways to solve the underdetermined BSS problem when the sources are sparse [5]–[7].

To simplify the derivation of the algorithm, in this letter we focus on the case $m = 2$. For this particular case, the kth column of the channel matrix can be parameterized as

$$\mathbf{h}(k) = \|\mathbf{h}(k)\| \left[\cos\theta_k \quad \sin\theta_k\right]^T, \qquad k = 0, \ldots, q-1 \tag{4}$$

where $\|\cdot\|$ denotes the L_2-norm of a vector, and θ_k is the angle indicating the direction of the kth column of the channel matrix. As an example, Fig. 1 shows a scatter plot of the observations for the filter in (13) ($m = 2$, $q = 3$) with $N = 10\,000$ and $p = 0.75$.

III. Algorithm Description

The proposed algorithm consists of three stages: identifying the angles of the q columns of $\mathbf{H}$, estimating their norms, and eliminating the ambiguity in the ordering of the columns.

A. Identifying the Angles

Since the input signals are highly sparse, in many occasions only one input sample will be nonzero. Assuming that for $n, \ldots, n-k, \ldots, n-q+1$, only $s(n-k)$ is nonzero, then $\mathbf{x}(n) = \mathbf{h}(k)s(n-k)$, and the output is collinear with the kth column of $\mathbf{H}$. Hence, we may obtain an estimate of the kth angle as

$$\hat{\theta}_k = \arctan\left(\frac{x_2(n)}{x_1(n)}\right). \tag{5}$$

Hence, as long as the sparsity factor is high, the angles can be estimated as the directions of maximum data density [6], [7]. Specifically, we use the method developed in [7] for BSS: construct a histogram, select the q bins with the greatest number of elements and estimate the angles as the sample mean of all the elements inside the selected bins.

B. Estimating the L_2-Norm of the Columns

The second step of the algorithm consists of estimating the L_2-norm of each column. Specifically, our objective is to obtain maximum-likelihood (ML) estimates of $\|\mathbf{h}(k)\|$, for $k = 0, \ldots, q-1$. The estimate for the kth column can be obtained using the set of N_k observations lying along the direction θ_k, which we denote as $[x_{1,k}(r), x_{2,k}(r)]$, for $r = 1, \ldots, N_k$. Once the angle is estimated, the observations can be expressed as

$$\begin{bmatrix} x_{1,k}(r) \\ x_{2,k}(r) \end{bmatrix} = \|\mathbf{h}(k)\| \begin{bmatrix} \cos\hat{\theta}_k \\ \sin\hat{\theta}_k \end{bmatrix} s_k(r), \qquad r=1, \ldots, N_k \tag{6}$$

where $s_k(r)$ is a sample driven from a zero-mean Gaussian distribution with unit variance. Therefore, $x_{1,k}(r)$ and $x_{2,k}(r)$ are also zero-mean Gaussian distributions with variance $\sigma_1^2 = \|\mathbf{h}(k)\|^2 \cos^2\hat{\theta}_k$ and $\sigma_2^2 = \|\mathbf{h}(k)\|^2 \sin^2\hat{\theta}_k$, respectively.

On the other hand, each observation has a joint probability density function that can be obtained through

$$f\left(x_{1,k}(r), x_{2,k}(r)\right) = f\left(x_{2,k}(r) | x_{1,k}(r)\right) f\left(x_{1,k}(r)\right). \tag{7}$$

In a noiseless situation and considering that the angles are known, $x_{2,k}(r) = x_{1,k}(r)\tan\hat{\theta}_k$ and then

$$f\left(x_{2,k}(r) | x_{1,k}(r)\right) = \delta\left(x_{2,k}(r) - x_{1,k}(r)\tan\hat{\theta}_k\right) \tag{8}$$

where $\delta(\cdot)$ is Dirac's delta. Therefore, estimating $\|\mathbf{h}_k\|^2$ amounts to estimating the variance of either $x_{1,k}(r)$ or $x_{2,k}(r)$. By grouping the N_k observations into vectors $\mathbf{x}_{1,k} = [x_{1,k}(1), \ldots, x_{1,k}(N_k)]^T$ and $\mathbf{x}_{2,k} = [x_{2,k}(1), \ldots, x_{2,k}(N_k)]^T$, the ML estimate of the L_2-norm of the kth column is given by [9]

$$\left\|\hat{\mathbf{h}}_k\right\| = \sqrt{\frac{\mathbf{x}_{1,k}^T \mathbf{x}_{1,k}}{N_k \cos^2\hat{\theta}_k}} \tag{9}$$

or alternatively

$$\left\|\hat{\mathbf{h}}_k\right\| = \sqrt{\frac{\mathbf{x}_{2,k}^T \mathbf{x}_{2,k}}{N_k \sin^2\hat{\theta}_k}}. \tag{10}$$

If the angles were known without errors, both estimates would be identical. However, we must admit some error in the estimate $\hat{\theta}_k = \theta_k + \Delta\theta_k$. Therefore, it is preferable to estimate the norm as

$$\left\|\hat{\mathbf{h}}_k\right\| = \sqrt{\frac{\mathbf{x}_{1,k}^T \mathbf{x}_{1,k} + \mathbf{x}_{2,k}^T \mathbf{x}_{2,k}}{N_k}}. \tag{11}$$

TABLE I
SUMMARY OF THE CHANNEL IDENTIFICATION ALGORITHM

1.- Identifying the Angles:
- 1.1- Estimate the angles for each output vector using (5).
- 1.2- Construct a histogram of angles.
- 1.3- Search for the n bins with the highest peaks of the histogram.
- 1.4- Obtain the n angles as the sample mean inside the selected bins.

2.- Estimating the scale factor:
- 2.1- Select the output samples collinear with each column.
- 2.2- Estimate $\|\mathbf{h}_k\|$ by applying (11).

3.- Sorting the columns of $\mathbf{H}$:
- 3.1- Find sequences of q consecutive outputs collinear with some column.
- 3.2- Obtain the most likely column ordering (i.e., the one that occurs most often).

C. Sorting the Columns of the Channel Matrix

Thus far we have solved the first two steps of the blind identification problem, i.e., we have identified the mixing matrix up to a global scale factor ambiguity and a permutation in the columns of $\mathbf{H}$. These two steps solve the BSS problem. However, in a deconvolution problem the columns of the mixing matrix must be ordered. This permutation may be eliminated by exploiting the strong temporal correlation that exists between consecutive input vectors.

The ordering method is based on the observation that a nonzero sample surrounded by $q-1$ zeros, i.e., a segment of the BG signal of the form

$$\underbrace{0\ldots0}_{q-1}, s(n), \underbrace{0\ldots0}_{q-1} \tag{12}$$

will be consecutively collinear with the q columns of the channel matrix $\mathbf{H}$. Hence, we can estimate the column ordering considering q consecutive output samples that are collinear with some column of $\mathbf{H}$ and setting the most likely column ordering as the one that appears most often. A summary of the overall algorithm is presented in Table I.

IV. SIMULATION RESULTS

To validate the performance of the algorithm, we have considered the identification of the following SIMO FIR channel

$$\begin{aligned}\mathbf{H}_1 &= \begin{bmatrix} \cos\left(\frac{\pi}{4}\right) & 0.3\cos\left(\frac{-7\pi}{12}\right) & 0.7\cos\left(\frac{2\pi}{9}\right) \\ \sin\left(\frac{\pi}{4}\right) & 0.3\sin\left(\frac{-7\pi}{12}\right) & 0.7\sin\left(\frac{2\pi}{9}\right)\end{bmatrix} \\ &= \begin{bmatrix} 0.707 & -0.077 & 0.536 \\ 0.707 & -0.289 & 0.450\end{bmatrix}.\end{aligned} \tag{13}$$

We carried out a Monte Carlo analysis for sparsity factors ranging from 0.5 to 0.95 and registers lengths from 500 to 5000 samples. The mse of the estimated SIMO channel was used as a figure of merit to evaluate the performance. For each sparsity factor and each register's length, the mse was evaluated by averaging the results of 1000 independent trials. As a comparison, we include the results obtained with the subspace method (SSM) described in [2]. The results are shown in Fig. 2. The proposed technique performs well for sparsity factors greater than 0.5, providing the best results in the range 0.6 to 0.8. For lower sparsity factors, the sources cannot be considered sparse, and the performance of the algorithm breaks down, whereas for sparsity factors larger than 0.8, the number of samples used to estimate the L_2-norm of the columns decreases, therefore increasing the variance of the final estimate.

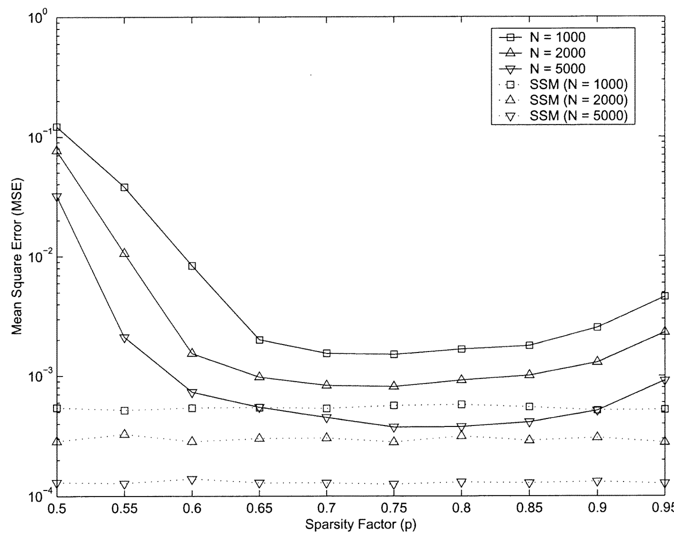

Fig. 2. MSE as a function of the sparsity factor and the number of samples.

Obviously, the SSM algorithm obtains better results, but we must remark the difference in computational cost between both procedures: the subspace-based technique requires an eigenvalue decomposition of an $mN \times mN$ autocorrelation matrix, thus requiring $O((mN)^3)$ floating-point operations (flops). For the proposed algorithm, the computational cost is mainly due to the angle estimation step, which only requires $O((m-1)N)$ flops.

V. CONCLUSION

In this letter, we have proposed a fast blind SIMO channel identification algorithm that exploits the sparsity of the input signals. The computational cost of the proposed algorithm is much lower than that of subspace-based methods, but maintains good performance, at least in the noiseless case. Future work will include extending the algorithm for higher oversampling ratios ($m > 2$) and studying its performance under noisy conditions, as well as with real data.

References

[1] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization of multipath channels," in *Proc. 25th Asilomar Conf. Acoustics, Speech and Signal Processing*, May 1995, pp. 1964–1967.

[2] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 516–525, Feb. 1995.

[3] K. Abed-Meraim, E. Moulines, and P. Loubaton, "Prediction error methods for second-order blind identification," *IEEE Trans. Signal Processing*, vol. 45, pp. 694–705, Mar. 1997.

[4] M. I. Gurelli and C. L. Nikias, "EVAM: An eigenvector-based algorithm for multichannel blind deconvolution of input colored signals," *IEEE Trans. Signal Processing*, vol. 43, pp. 134–149, Jan. 1995.

[5] T. W. Lee, M. S. Lewicki, M. Girolami, and T. J. Sejnowski, "Blind source separation of more sources than mixtures using overcomplete representations," *IEEE Signal Processing Lett.*, vol. 6, pp. 87–90, Apr. 1999.

[6] P. Bofill and M. Zibulevsky, "Underdetermined blind source separation using sparse representations," *Signal Processing*, vol. 81, no. 11, pp. 2353–2362, Nov. 2001.

[7] D. Erdogmus, L. Vielva, and J. C. Principe, "Nonparametric estimation and tracking of the mixing matrix for underdetermined blind source separation," in *Proc. 3rd Int. Conf. Independent Component Analysis and Signal Separation*, San Diego, CA, Dec. 9–13, 2001, pp. 189–193.

[8] J. M. Mendel, *Maximum Likelihood Deconvolution*. New York: Springer-Verlag, 1990.

[9] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.