BLIND DECONVOLUTION OF MULTI-INPUT SINGLE-OUTPUT SYSTEMS WITH BINARY SOURCES

Konstantinos I. Diamantaras

Department of Informatics
TEI of Thessaloniki
Sindos 54101, Greece
kdiamant@it.teithe.gr

Theophilos Papadimitriou

Department of Int. Economic Relat. and Devel.,
Democritus University of Thrace
Komotini 69100, Greece
papadimi@uom.gr

ABSTRACT

We present two novel methods for the blind extraction of multiple binary signals from a single observation generated from linear convolution and superposition. Both methods are based on two features: (a) the constellation structure of the successor set for each output value, and (b) the concept of system deflation, ie. the recursive shortening of the channel until it is reduced into an instantaneous mixture. The two methods differ in complexity as the latter requires significantly smaller amounts of data. Both methods are fast for small problem sizes but their complexities increase exponentially as the number of inputs and/or channel length increases. We present the complete analytical solution to the noiseless case and we thoroughly treat the related mathematical tools and concepts. The noisy case is briefly discussed as the complete treatment would exceed the scope of the present paper.

1. INTRODUCTION

The Blind Deconvolution of a MIMO system refers to the extraction of the unknown inputs from the observed outputs, without knowledge of the system itself. This problem is also referred to as Blind Signal Separation (BSS) from linear convolutive mixtures. This problem is very important, for example, in wireless communications, where n transmitted signals corrupted by intersymbol interference (ISI), multiuser interference (MUI), and noise are received at m antennas. In general, sufficient number of receivers ($m \geq n$) guarantee the deconvolution of the signals using various techniques [1, 2]. On the other hand, in the case of multiuser digital subscriber lines (DSL) and asychronous DSL (ADSL), n sources are mixed in one signal and transmitted through a telephone line. In typical telephone communication standards, several lines are grouped in a cable bundle, introducing crosstalk effects. This system is underdetermined and is rarely studied. It can be modeled using the Multi-Input Single-Output (MISO) model.

In this paper we treat MISO systems with binary inputs. The blind deconvolution of a MISO system with finite alphabet signals is discussed in [3]. The approach is algebraic and converts the nonlinear system into a linear one using the finite alphabet property. The application of a MIMO technique to a MISO system was proposed in [4]. The method behaves almost as good as in the MIMO case. In [5], a semi-blind technique is presented, in which training sequences of the system are available and they can be used to acquire channel estimates. Aldana *et al.* in [6] treats the input symbols as discrete random variables in a stochastic likelihood criterion. The system is solved by applying the Expectation-Maximization algorithm in the frequency domain.

The approach presented here is essentially geometric. It is not related to the second or higher order statistics of the signals. Instead, we exploit the constellation structure of the successors to each output value. Our model is described by the following equation:

$$x(k) = \sum_{l=0}^{L-1} \mathbf{a}_l^T \mathbf{s}(k-l) + e(k), \quad k = 1, ..., K \qquad (1)$$

where $\mathbf{a}_l$ for $l = 0, ..., L-1$ are a set of unknown real n-dimensional mixing vectors, $\mathbf{s}(k) = [s_1(k) \cdots, s_n(k)]^T$ is the n-dimensional source signal, and $e(k)$ corresponds to additive gaussian noise. The n independent binary sources take discrete values from the set $\{-1, +1\}$. The observations of the mixtures are real-valued scalars. We start by ignoring the noise in Section 2 so that the underlying mathematical ideas are clearly introduced. Based on these ideas two alternative methods are proposed in Sections 2.1 and 2.2. The more realistic noisy model is briefly discussed in Section 3. A more elaborate treatment of the noise is reserved for the full paper. The present work is an extension of [7] where the problem for $n = 1$, $L = 1$ has been treated.

2. THE NOISELESS CASE

Two methods will be presented in the noseless case. Both exploit the constellation structure of the successor values in order to achieve recursive *system deflation*. The second method is more efficient compared to the first one in terms of required data lengths.

2.1. First Method

Each source sample $\mathbf{s}(k)$ is a vector composed of n binary elements and, thus, it can take exactly 2^n values denoted by

$$\begin{aligned}
\mathbf{v}_1 &= [-1,-1,-1,\cdots,-1,-1,-1]^T \\
\mathbf{v}_2 &= [-1,-1,-1,\cdots,-1,-1,+1]^T \\
&\vdots \\
\mathbf{v}_{2^n-1} &= [+1,+1,+1,\cdots,+1,+1,-1]^T \\
\mathbf{v}_{2^n} &= [+1,+1,+1,\cdots,+1,+1,+1]^T \qquad (2)
\end{aligned}$$

Moreover, each observation $x(k)$ is generated by the linear combination of L source vectors. Consequently, the observation space $\mathcal{X}$ is a discrete set consisting of, at most, 2^M elements, ie. $|\mathcal{X}| \leq 2^M$, $M = nL$. We will have $|\mathcal{X}| < 2^M$ if and only if there exist two separate L-tuples $\{\mathbf{b}_0,\cdots,\mathbf{b}_{L-1}\}$ and $\{\mathbf{b}'_0,\cdots,\mathbf{b}'_{L-1}\}$, of binary vectors such that $\sum_{l=0}^{L-1}\mathbf{a}_l^T\mathbf{b}_l = \sum_{l=0}^{L-1}\mathbf{a}_l^T\mathbf{b}'_l$. We avoid this situation by making the following assumption:

Assumption 1 *For every observation value $r = x(k) \in \mathcal{X}$, some k, there exists a unique L-tuple $\{\mathbf{b}_0, \mathbf{b}_1, \cdots, \mathbf{b}_{L-1}\}$ of source vectors $\{\mathbf{s}(k), \mathbf{s}(k-1), \cdots, \mathbf{s}(k-L+1)\}$ that generates r according to Eq.* (1). ■

From Assumption 1 it follows that the cardinality of $\mathcal{X}$ is exactly 2^M. In the following we will use, when necessary, the notation $\{\mathbf{b}_0(r), \mathbf{b}_1(r), \cdots, \mathbf{b}_{L-1}(r)\}$ to denote the unique L-tuple of source vectors that generates the observation $r \in \mathcal{X}$.

Let us now take any observation value $r \in \mathcal{X}$ and let $x(k) = r$, for some k. The successor observation $x(k+1)$ is:

$$\begin{aligned}
x(k+1) &= \mathbf{a}_0^T\mathbf{s}(k_0+1) + \sum_{l=1}^{L-1}\mathbf{a}_l^T\mathbf{s}(k-l+1) \\
x(k+1) &= \mathbf{a}_0^T\mathbf{s}(k_0+1) + \sum_{l=1}^{L-1}\mathbf{a}_l^T\mathbf{b}_{l-1}(r) \qquad (3)
\end{aligned}$$

since, by definition, for $l = 0,\cdots,L-1$,

$$\mathbf{b}_l(r) = \mathbf{s}(k-l).$$

The first term on the right-hand-side of Eq. (3) depends on the binary vector $\mathbf{s}(k+1)$ which can take 2^n distinct values, $\mathbf{v}_1$, ..., $\mathbf{v}_{2^n}$. The second term depends only on r. Therefore, for a fixed r, there exist 2^n possible values for $x(k+1)$ called *successors* of $x(k)$ and denoted by $\sigma_i(r)$, $i = 1,\cdots,2^n$:

$$\sigma_i(r) = \mathbf{a}_0^T\mathbf{v}_i + \sum_{l=1}^{L-1}\mathbf{a}_l^T\mathbf{b}_{l-1}(r) \qquad i = 1,\cdots,2^n (4)$$

Let us now compute the mean $\mu(r)$ of the 2^n successors, noting that $\sum_{j=1}^{2^n}\mathbf{v}_j = 0$, to obtain:

$$\begin{aligned}
\mu(r) &= \frac{1}{2^n}\sum_{i=1}^{2^n}\sigma_i(r) \\
&= \frac{1}{2^n}\left(\mathbf{a}_0^T\sum_{j=1}^{2^n}\mathbf{v}_j + 2^n x^{(2)}(k)\right) \\
&= x^{(2)}(k) \qquad (5)
\end{aligned}$$

where

$$\begin{aligned}
x^{(2)}(k) &= \sum_{l=1}^{L-1}\mathbf{a}_l^T\mathbf{b}_{l-1}(r) \\
x^{(2)}(k) &= \sum_{l=1}^{L-1}\mathbf{a}_l^T\mathbf{s}(k-l+1) \qquad (6)
\end{aligned}$$

Eq. (6) representes a shortened MISO system compared to the original system of Eq. (1). It is clear that the new MISO system has the same taps as the original one except for the absence of $\mathbf{a}_0$ and the input time-shift. Of course the length of the new system is $L-1$, i.e. one less than the initial length L. Based on the above analysis the method for reducing the system length is described by the following steps:

- Step 1. For every k locate the set of points $x(k_j)$, $j = 1,2,\cdots$, for which $x(k_j) = r = x(k)$.
- Step 2. Find the set $\Sigma(r) = \{x(k_j+1);\ j = 1,2,\cdots\}$ of the successors of $x(k)$
- Step 3. Identify 2^n distinct values $\sigma_i(r)$ in $\Sigma(r)$
- Step 4. Compute the mean $\mu(r) = 1/2^n\sum_i\sigma_i(r)$
- Step 5. Replace $x(k_j)$ by $\mu(r)$, for all j.

This procedure will be called *filter deflation* or *system deflation*. Clearly, it is essential for the present method that all observation/successor pairs $[r, \sigma_i(r)]$, $i = 1,\cdots,2^n$, will appear, at least once, in the output sequence x. Thus, we make the following assumption

Assumption 2 *For any $r \in \mathcal{X}$, there are at least 2^n time indices $k_1, k_2, \cdots, k_{2^n} \in \{1,2,\cdots,K\}$ such that $x(k_i) = r$ and $x(k_i+1) = \sigma_i(r)$, where $i = 1,\cdots,2^n$.* ■

The above deflation method can be recursively applied $L-1$ times until the system is reduced into:

$$x^{(L)}(k) = \mathbf{a}_{L-1}^T \mathbf{s}(k-L+1) \quad (7)$$

The estimation of the sources $\mathbf{s}$ from Eq. (7) has been treated elsewhere (see [8, 9]).

2.2. Second Method

The main disadvantage of the method presented in section 2.1, comes as a result of the second assumption, where it is stated that every possible pair of observations must exist in the dataset. As the complexity of the MISO system increases (more sources and more filter taps), the validity of the second assumption requires exponentially larger observation datasets. In order to treat cases where large datasets are not available, we propose a second method based on the same principles. This second method also requires the validity of Assumption 1. However, instead of Assumption 2, a less restrictive one will be introduced next:

Assumption 3 *For* **only one** $r_0 \in \mathcal{X}$*, there exist at least* 2^n k_i*,* $i = 1, \cdots, 2^n \in \{1, 2, \cdots, K\}$ *such that* $x(k_i) = r_0$*,* $x(k_i+1) = \sigma_i(r_0)$*,* $i = 1, \cdots, 2^n$*. In addition to that, every possible value of* $\mathcal{X}$ *exists at least once in the dataset.* ■

Equiped with Assumption 3, instead of Assumption 2, we can still compute the following values

$$c_i = \mathbf{a}_0^T \mathbf{v}_i = \sigma_i(r_0) - \mu(r_0) \qquad i = 1, \cdots, 2^n \quad (8)$$

We shall refer to the set $C = \{c_i;\ i = 1, \cdots, 2^n\}$ as the *successor constellation set* of system (1). Now, for every observation value $r = x(k) \in \mathcal{X}$ we have

$$r = \mathbf{a}_0^T \mathbf{v}_i + \sum_{l=1}^{L-1} \mathbf{a}_l^T \mathbf{b}_l(r) \quad (9)$$

$$= c_i + \sum_{l=1}^{L-1} \mathbf{a}_l^T \mathbf{b}_l(r), \qquad \text{some } i \quad (10)$$

Furthermore, due to the symmetry of the constellation set, there exists a “dual” observation value $r^d \in \mathcal{X}$ such that

$$r^d = -c_i + \sum_{l=1}^{L-1} \mathbf{a}_l^T \mathbf{b}_l(r) \quad (11)$$

so

$$r^d = r - 2c_i \quad (12)$$

The mean of the two values is

$$\mu_2(r) = (r + r^d)/2 = \sum_{l=1}^{L-1} \mathbf{a}_l^T \mathbf{b}_l(r) \quad (13)$$

Note that $\mathbf{b}_l(r) = \mathbf{s}(k-l)$. Therefore, if we replace $x(k)$ by $\mu_2(r)$ we obtain a new, shortened MISO system, similar (although not identical) to (6)

$$\tilde{x}^{(2)}(k) = \sum_{l=1}^{L-1} \mathbf{a}_l^T \mathbf{s}(k-l) \quad (14)$$

The problem now is to identify the dual observation r^d. The next assumption leads to the solution.

Assumption 4 *For every observation* $r \in \mathcal{X}$*, there exists a unique* $j \in \{1, \cdots, 2^n\}$ *such that* $r - 2c_j \in \mathcal{X}$*.* ■

Using Assumption 4 the dual value r^d can be found by testing all $r - 2c_j$, $j = 1, \cdots, 2^n$, for membership in the observation space $\mathcal{X}$.

Summarizing the above results, our second method for obtaining the deflated system (14) is described below:

Step 1. Locate an observation value r_0 for which 2^n distinct successors $\sigma_i(r_0)$, $i = 1, \cdots, 2^n$, exist in the dataset

Step 2. Compute the successor constellation set C according to (8)

Step 3. For every observation $r = x(k)$ find the (unique) value j for which $r - 2c_j \in \mathcal{X}$. Call $r^d = r - 2c_j$.

Step 4. Replace $x(k)$ by $(r + r^d)/2$

Again, the $L-1$ times repetition of this algorithm will reduce the system into a memoryless one

$$\tilde{x}^{(L)}(k) = \mathbf{a}_{L-1}^T \mathbf{s}(k) \quad (15)$$

which can be treated as described in section 2.1.

2.3. Comparing the two methods

The main advantage of the second method over the first one is that it requires quite smaller datasets. In order to estimate this advantage, we randomly created large datasets and we identified the minimum dataset which satisfies the assumptions for each method. We tested 100 datasets for different MISO systems and we kept the worst case (the maximum dataset length). Results of the conducted experiments are presented in Table 1. It can be observed that the gain increases as the system becomes more complex. For SISO Systems the gain is around 50%, while for more complex cases ($L > 2$ and $n > 2$) the gain increases upto 90%. It is noteworthy that in a Pentium III 2.6 GHz processor, the MATLAB implementation of method 2 with $K = 100,000$ samples, $n = 3$, and $L = 3$, runs in less than 5 seconds.

Table 1. Dataset sizes satisfying the assumptions of methods 1 and 2, and relative dataset-size reduction.

n	L	Method 1	Method 2	Gain (%)
1	2	64	33	48.44
1	3	133	72	45.86
1	4	396	176	55.56
2	1	98	39	60.20
2	2	658	108	83.59
2	3	3228	685	78.78
2	4	11801	2678	77.31
3	1	499	148	70.34
3	2	6888	657	90.46
3	3	49994	6349	87.30
4	1	2657	611	77.00
4	2	61605	6206	89.93

3. THE NOISY CASE

The above methods can be extended to noisy systems although detailed discussion of this case is impossible due to lack of space. Here we shall only outline a simple case where the observation is infected with low-level noise. The general model that represents an observation $x(k)$ is:

$$x(k) = r + e(k), \quad k = 1, ..., K \tag{16}$$

where r is a noiseless *prototype value*, ie. a member of the observation space $\mathcal{X}$. Remember that there exist 2^{nL} such prototypes. Each prototype r is associated with a class $\mathcal{I}_r$ of observations $x(k) \in \mathcal{I}_r$, that are noisy versions of r according to (16). If the noise level is sufficiently low then $\mathcal{I}_r$ is defined as $\{x(k);\ r = \arg\min_{\rho} |x(k) - \rho|\}$. Thus every noisy observation $x(k)$ would be successfully clustered in the appropriate class. In order to estimate the prototypes r a clustering process can be used. Then every $x(k) \in \mathcal{I}_r$ is replaced by the estimate of the cluster center $\hat{r}$ and the Blind Deconvolution method proceeds as in the noiseless case. We succesfully tested the above approach using a clustering scheme based on the properties of the statistical median and the k-means clustering. However, higher noise levels render this approach useless since a large number of output samples are misclassified. Due to lack of space detailed treatement of this case is postponed for future work.

4. CONCLUSION

A novel approach for the blind deconvolution of MISO systems with binary sources was presented in this paper. We introduce the concept of the successor constellation set which is instrumental for the development of a system deflation process. Recursive application of system deflation leads to a linear memoryless system. We present two methods based on these ideas. The first method requires large datasets in order to successfully estimate the sources, while the second one is less data-demanding yet more complex. The paper describes in detail the noise-free case and gives only a brief outline of a simple noisy scenario. The complete treatment of the noisy case is reserved for a future publication.

5. REFERENCES

[1] S. Talwar, M. Viberg, and A. Paulraj, "Blind Separation of Synchronous co-channel Digital Signals using an Antenna Array - Part I: Algorithms," *IEEE Trans. Signal Process.*, vol. 44, pp. 1184–1197, May 1996.

[2] K. Abed-Meraim, P. Loubaton, and E. Moulines, "A Subspace Algorithm for certain Blind Identification Problems," *IEEE Trans. Inform. Theory*, pp. 499–511, March 1997.

[3] D. Yellin and B. Porat, "Blind Identification of FIR systems excited by discrete alphabet inputs," *IEEE Trans. Signal Process.*, vol. 41, no. 3, pp. 1331–1339, March 1993.

[4] N. Al-Dhahir and A. Sayed, "The finite-length Multi-Input Multi-Output MMSE-DFE," *IEEE Trans. Signal Process.*, vol. 48, pp. 2921–2936, October 2000.

[5] Y. Xie and C. Georghiades, "Two EM-type Channel Estimation Algorithms for OFDM with Transmitter Diversity," *IEEE Trans. on Commun.*, vol. 52, pp. 106–115, January 2003.

[6] C. Aldana, E. de Carvalho, and J. Cioffi, "Channel Estimation for Multicarrier Multiple Input Single Output Systems using the EM Algorithm," *IEEE Trans. Signal Process.*, vol. 51, pp. 3280–3292, December 2003.

[7] K. Diamantaras and T. Papadimitriou, "Blind Deconvolution of SISO Systems with Binary Source based on Recursive Channel Shortening," in *Proc. 5th Int. Symposium on Independent Component Analysis and Blind Signal Separation*, Granada, Spain, September 2004.

[8] K. I. Diamantaras, "Blind Separation of Multiple Binary Sources using a Single Linear Mixture," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP-2000)*, vol. V, Istanbul, Turkey, June 2000, pp. 2889–2892.

[9] K. I. Diamantaras and E. Chassioti, "Blind separation of n binary sources from on observation: A deterministic approach," in *Proc. Second Int. Workshop on ICA and BSS*, Helsinki, Finland, June 2000, pp. 93–98.