DIRECT BLIND MULTICHANNEL IDENTIFICATION

M. Frikel *, W. Utschick and J. Nossek
Technical University of Munich
Institute for Circuit Theory and Signal Processing
Arcisstr. 21, D-80290 Munich, Germany
e-mail: mifr@nws.e-technik.tu-muenchen.de

ABSTRACT

In this paper, different techniques for the estimation of the signal parameters and/or the channel coefficients for single-input/multiple-output systems, are presented. These methods are based, either on the use of a small part of observations or on the minimization of a quadratic form with quadratic constraint. Simulations have been established and these techniques are compared to the classical approaches.

1 INTRODUCTION

Most standard approaches in array signal processing make no use of any available information about the signal structure. However, many man-made signals have a rich structure that can be used to improve the estimator performance. In digital communication applications, the transmitted signals are often cyclostationary, which implies that their autocorrelation functions are periodic. Different approaches using signal structure is based on high-order statistics. As is well-known, all information about a Gaussian signal is conveyed in the first and second order moments.

Recently, several techniques are presented for blind source separation and channel identification using second order statistics. Among the existing second-order statistics based methods, the subspace method shown in [1] appears to be very robust to noise. The concept behind the subspace method is also versatile, which not only has proved to be useful for many array processing problems. It has already been shown that blind identification is feasible based on spatial covariance matrices. These matrices show a simple structure that allows straightforward blind identification procedures based on eigendecomposition. When the source signals are temporally correlated but still mutually independent, it is possible to base source separation on this property. Also, if the source signals have different spectra, then it is easy to identify the transfer matrix. If sources are temporally correlated with different spectra, then separation can be achieved using only second-order statistics. In this paper, we present some techniques to estimate, "blindly" and "directly", the channel matrix and the signal parameters. The first technique is based on the minimization of quadratic form with a quadratic constraint, this method uses only a part of received data. The second technique is an improved approach of the technique presented in [3]. Finally, we compare these methods to existing techniques.

*This work is supported by Alexander von Humboldt-Stiftung

2 PROBLEM FORMULATION

Let the following notations: $\mathbf{A}^T$, $\mathbf{A}^+$, $\mathbf{A}^*$ are, respectively, the transpose, the transpose conjugate and the conjugate of the matrix $\mathbf{A}$. $diag(.)$ is the diagonal matrix. $\mathbf{I}$ is the identity matrix, and $E[.]$ is the statistical expectation. In the mathematical formulas: matrices (capital) "$\mathbf{A}$" and vectors "$\mathbf{a}$" are in boldface type. The scalars are represented by normal character "a". N, P and T are the number of sensors, number of sources and the number of realizations, respectively.

Consider a generic P-input/N-output causal discrete-time noisy linear time-invariant digital (MIMO) system describe by the convolution equation:

$$\mathbf{r}(k) = \sum_{p=1}^{P} \sum_{m=0}^{M_p-1} \mathbf{h}_p(m) s_p(k-m) + \mathbf{n}(k), \qquad (1)$$

where $\mathbf{r}(k)$ is an N-dimensional vector of system outputs, $\{\mathbf{h}_p(l) : l = 0, 1, ..., L_p - 1\}$ is the finite impulse response (FIR) associated to the pth user's scalar input signal $s_p(k)$, and $\mathbf{n}(k)$ denotes the additive noise. The precedent equation can be rewritten as:

$$\mathbf{r}(k) = \sum_{p=1}^{P} \mathbf{H}_p \mathbf{s}_p(k) + \mathbf{n}(k), \qquad (2)$$

$$\mathbf{r}(k) = \mathbf{H}\mathbf{s}(k) + \mathbf{n}(k) = \mathbf{x}(k) + \mathbf{n}(k), \qquad (3)$$

where $\mathbf{x}(k)$ is the LN noise-free output vector, $\mathbf{s}(k)$ is the $(N+M)$ input vector (source signals), $\mathbf{H}_p$ is defined as $LN \times (N+M)$ matrix, $\mathbf{H}_p = [\mathbf{h}_p(0) \quad ... \quad \mathbf{h}_p(L_p-1)]$, and $\mathbf{H} = [\mathbf{H}_1 \quad ... \quad \mathbf{H}_P]$ is the $LN \times (N+M)$ channel convolution matrix. $L = L_1 + L_2 + ... + L_P$ is the overall order of the system. The vector $\mathbf{s}(k)$ is obtained by stacking the P vectors $\mathbf{s}_p(k) =$

$[s_p(k)...s_p(k-L_p+1)]^T$. The input signals $s_p(k)$ are taken from a finite modulation-dependent alphabet α. In SDMA communications, the precedent equation models the complex baseband transfer function between the information symbols $s_p(k)$ emitted by the sources and the oversampled array outputs [4]. In this contribution, we assume the case of one user (P=1).

Let $\mathbf{r}_l^{(i)} = \left[r_l^{(i)}, \cdots, r_{n-N+1}^{(i)}\right]^T$,

$\mathbf{n}_l^{(i)} = \left[n_l^{(i)}, \cdots, n_{n-N+1}^{(i)}\right]^T$,

and $\mathbf{s}_l = [s_l, \cdots, s_{n-N-M+1}]^T$.

$$\mathbf{r}_l^{(i)} = \mathcal{H}_l^{(i)}\mathbf{s}_l + \mathbf{n}_l^{(i)}, \quad i = 1, \cdots L$$

$$\begin{pmatrix} \mathbf{r}_l^{(1)} \\ \vdots \\ \mathbf{r}_l^{(L)} \end{pmatrix} = \begin{pmatrix} \mathcal{H}_l^{(1)} \\ \vdots \\ \mathcal{H}_l^{(L)} \end{pmatrix} S_l + \begin{pmatrix} \mathbf{n}_l^{(1)} \\ \vdots \\ \mathbf{n}_l^{(L)} \end{pmatrix}$$

$$\mathbf{r}_l = \mathcal{H}_l\mathbf{s}_l + \mathbf{n}_l, \tag{4}$$

with, $\mathcal{H}_l$ is the channel matrix. To solve this problem we assume the following conditions:

A1: The matrix $\mathcal{H}$ is full column rank.

A2: The signal $\mathbf{s}_l$ is statistically independent.

A3: The noise $\mathbf{n}_l$ are Gaussian random vectors, independent of the signal and independent of each other with zero mean and covariance matrix $\sigma^2\mathbf{I}$.

It follows from these assumptions that the $(LN \times LN)$ covariance matrix of the observation vector $\mathbf{r}(k)$ is given by: $\mathbf{\Gamma} = E\left[\mathbf{r}_l\mathbf{r}_l^+\right]$. The identification is based on the $LN \times LN$ received covariance matrix $\mathbf{\Gamma} = \mathcal{H}_l\mathbf{\Gamma}_s\mathcal{H}_l^+ + \mathbf{\Gamma}_n$, Let $\mathbf{\Gamma}_n = \sigma^2\mathbf{I}$ and $\mathbf{\Gamma}_s = E\left[\mathbf{s}_l\mathbf{s}_l^+\right]$ are the noise and signal covariance matrices, respectively. Below, a "special" partition of the received data is used in order to estimate the channel parameters.

3 MODELING OF THE RECEIVED SIGNAL and BLIND CHANNEL IDENTIFICATION

It's assumed that the matrix $\mathcal{H}$ is of full rank $(N+M)$, that is, $(N+M)$ rows of $\mathcal{H}$ are linearly independent and hence the other rows can be expressed as a linear combination of these $(N+M)$ rows. Hereafter, the first $(N+M)$ rows are assumed to be linearly independent and the received signal vector $\mathbf{r}$ is partioned as follows:

$$\begin{bmatrix} \mathbf{r}_1 \\ - \\ \mathbf{r}_2 \end{bmatrix} = \begin{bmatrix} \mathcal{X} \\ - \\ \mathcal{Y} \end{bmatrix} \mathbf{s} + \begin{bmatrix} \mathbf{n}_1 \\ - \\ \mathbf{n}_2 \end{bmatrix}, \tag{5}$$

where $\mathcal{X} \in \mathcal{C}^{(N+M)\times(N+M)}$ and $\mathcal{Y} \in \mathcal{C}^{LN\times(LN-(N+M))}$. The linear combiner or the operator $\Pi \in \mathcal{C}^{(N+M)\times LN-(N+M)}$ is defined as:

$$\Pi^+\mathcal{X} = \mathcal{Y}, \tag{6}$$

$$or, [\,\Pi^+ \mid -\mathbf{I}_{LN-(N+M)}\,] \begin{bmatrix} \mathcal{X} \\ - \\ \mathcal{Y} \end{bmatrix} = \mathbf{Q}^+\mathcal{H} = \mathbf{0},$$

where $\mathbf{I}_{LN-(N+M)}$ and $\mathbf{0}$ are the identity and the null matrices, respectively, and $\mathbf{Q} \in \mathcal{C}^{LN-(N+M)\times(N+M)}$. The precedent equation implies that the vectors of $\mathcal{H}$ are orthogonal to the columns of $\mathbf{Q}$. This means that the subspace spanned by the columns of the matrix $\mathbf{Q}$, span$\{\mathbf{Q}\}$, is included in span$\{\mathbf{V}_n\}$, where $\mathbf{V}_n$ is the noise subspace, i.e., the eigenvectors associated with the smallest eigenvalues of the data covariance matrix $\mathbf{\Gamma}$. Now, since $\mathbf{Q}$ contains the block $\mathbf{I}_{LN-(N+M)}$, its $LN-(N+M)$ columns are linearly independent, therefore, span$\{\mathbf{Q}\}$=span$\{\mathbf{V}_n\}$. It follows that the linear combiner defines the noise subspace. Note that, in contrast to the basis defined by $\mathbf{V}_n$, the basis defined by the columns of $\mathbf{Q}$ is not orthonormal. However, the result of applying Householder transforms to $\mathbf{Q}$ will enable us to find an orthonormal basis for $\mathbf{Q}$ and an orthonormal basis for its orthonormal complement, which is the signal subspace, $\mathbf{V}_s$.

3.1 Blind subspace method

Here, we recall the blind subspace approach [1], this method yields an estimate $\widehat{\mathcal{H}}$ of $\mathcal{H}$ by solving the equation $\mathbf{U}_n^+\widehat{\mathcal{H}} = \mathbf{0}$, in a least square sense (where $\widehat{\mathcal{H}}$ is subject to the same structure as $\mathcal{H}$). This estimate is uniquely (up to a constant scalar) equal to $\mathcal{H}$. From [1], we have, $\mathbf{U}_n^+\mathcal{H} = \mathbf{h}^+\mathcal{U}_n = \mathbf{0}$, where $\mathcal{U}_n$ is the $L(M+1)\times(N+M)$ matrix obtained by stacking the L filtering matrices $\mathcal{U}_n^{(l)}$: $\mathcal{U}_n = \left[\mathcal{U}_n^{(0)T}...\mathcal{U}_n^{(L-1)T}\right]^T$, where,

$$\mathcal{U}_n^{(l)} = \begin{pmatrix} u_1^{(l)} & u_2^{(l)} & \cdots & u_N^{(l)} & \cdots & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots & \ddots & \cdots & \vdots \\ 0 & \cdots & 0 & u_1^{(l)} & u_2^{(l)} & \cdots & u_N^{(l)} \end{pmatrix}.$$

The solution derived in [1] relies are the following optimization procedure:

$$\widehat{\mathbf{h}}_{SS} = \arg \min_{||\mathbf{h}||=1} \mathbf{h}^+\mathcal{U}\mathbf{h}, \tag{7}$$

with,

$$\mathcal{U} = \sum_{i=1}^{LN-M-N-1} \mathcal{U}_n^{(i)}\mathcal{U}_n^{(i)+}.$$

The solution of the precedent quadratic form under unit-norm constraint, is reached for $\widehat{\mathbf{h}}_{SS} = \widehat{\alpha}\widehat{\mathbf{v}}_N$, where $\widehat{\mathbf{v}}_N$ is the unit-norm eigenvector associated with the smallest eigenvalue of the matrix $\mathcal{U}$, for uniqueness, it is assumed that the first component of $\widehat{\mathbf{v}}_N$ is positive. And $\widehat{\alpha}$ is the missing scalar factor, representing the norm of $\mathbf{h}$ [6]: $\widehat{\alpha} =\mid tr\left(\frac{1}{N}(\widehat{\mathbf{\Gamma}} - \widehat{\sigma}^2\mathbf{I})\right)\mid^{\frac{1}{2}}$, $\widehat{\sigma}^2$ is the noise power.

3.2 Fast blind linear operator's method

The classical channel subspace method is based on the eigendecomposition of the data covariance matrix. However, this operation is too expensive in term of computational load. In some real time applications, the use of this method is difficult. In order to avoid this eigendecomposition, we suggest to use the linear operator defined above to estimate the noise subspace [4]. As, we showed the resulting noise subspace of the linear operator is equivalent to the eigenstructure noise subspace. Indeed, the estimation of the data covariance matrix and the noise subspace need $K(LN)^2 + (LN)^3$ using classical eigendecomposition, and $(LN)^2(K+N+M) + LN(N+M)^2$ for the linear operator. The fast blind channel identification [4] is based on the equation (7). The optimization criterion, when the linear operator is used, takes the form,

$$\widehat{\mathbf{h}}_{lin} = \arg \min_{||\mathbf{h}||=1} \mathbf{h}^+ \mathcal{Q}\mathbf{h}, \tag{8}$$

with,

$$\mathcal{Q} = \sum_{i=1}^{LN-M-N-1} \mathcal{Q}^{(i)} \mathcal{Q}^{(i)+},$$

$\mathcal{Q}^{(i)}$ is the $L(M+1) \times (M+N)$ filtering matrix associated with ith column $\mathbf{q}^{(i)}$ of $\mathbf{Q}$ [4]. The advantage is the non-eigendecomposition of the received covariance matrix. We note here that the channel parameter vector is estimated via a quadratic optimization whose solution can be obtained in closed form by using eigendecomposition. For the resolution of the precedent equation, many solutions had presented in [1]. We consider the minimization subject to a quadratic constraint. Therefore, minimize $\widehat{\mathbf{h}}$ subject to $\| \mathbf{h} \| = 1$; the solution is the unit-norm eigenvector associated to the smallest eigenvalue of $\mathcal{Q}$.

4 DIRECT BLIND MULTICHANNEL IDENTIFICATION

In this section, we propose to use a linear combination of the rows of the channel matrix. Indeed, it assumed that the matrix channel $\mathcal{H}$ is of full rank $(N+M)$, that is, $(N+M)$ rows of $\mathcal{H}$ are linearly independent and hence the other rows can be expressed as a linear combination of these $(N+M)$ rows. From the partition of the channel matrix $\mathcal{H} = \begin{bmatrix} \mathcal{X} \\ - \\ \mathcal{Y} \end{bmatrix}$, where $\mathcal{X}$ is a square matrix $(N+M) \times (N+M)$ and $\mathcal{Y}$ is a matrix of dimension $(LN-M-N) \times (N+M)$, we, only, estimate the channel matrix $\mathcal{X}$ of dimension $(N+M) \times (N+M)$. First, we consider the case of noise free:

$$\mathbf{x}_1(k) = \mathcal{X}\mathbf{s}(k),$$

where $\mathbf{x}_1(k)$ is a part of the vector observation $\mathbf{x}(k)$ of dimension $(N+M) \times 1$. The covariance matrix of $\mathbf{x}_1(k)$ is estimatated by $\mathbf{\Gamma}_{11} = E[\mathbf{x}_1(k)\mathbf{x}_1^+(k)]$, then $\mathbf{\Gamma}_{11} = \mathcal{X}\mathbf{\Gamma}_s\mathcal{X}^+$.

From the follwing partitions of the data covariances matrices,

$$\mathbf{\Gamma}(0) = \begin{bmatrix} \mathbf{\Gamma}_{11}(0) & | & \mathbf{\Gamma}_{12}(0) \\ \mathbf{\Gamma}_{21}(0) & | & \mathbf{\Gamma}_{22}(0) \end{bmatrix} = \begin{bmatrix} \mathcal{X}\mathcal{X}^+ & | & \mathcal{X}\mathcal{Y}^+ \\ \mathcal{Y}\mathcal{X}^+ & | & \mathcal{Y}\mathcal{Y}^+ \end{bmatrix},$$

and the partition of $\mathbf{\Gamma}(1)$,

$$\mathbf{\Gamma}(1) = \begin{bmatrix} \mathbf{\Gamma}_{11}(1) & | & \mathbf{\Gamma}_{12}(1) \\ \mathbf{\Gamma}_{21}(1) & | & \mathbf{\Gamma}_{22}(1) \end{bmatrix} = \begin{bmatrix} \mathcal{X}\mathbf{J}_1\mathcal{X}^+ & | & \mathcal{X}\mathbf{J}_1\mathcal{Y}^+ \\ \mathcal{Y}\mathbf{J}_1\mathcal{X}^+ & | & \mathcal{Y}\mathbf{J}_1\mathcal{Y}^+ \end{bmatrix},$$

we have, $\mathbf{\Gamma}_{11}(0) = \mathcal{X}\mathcal{X}^+$ and $\mathbf{\Gamma}_{11}(1) = \mathcal{X}\mathbf{J}_1\mathcal{X}^+$, with $\mathbf{J}_1$ is the shift matrix.

It is easy to see that:

$$\mathbf{x}_i^+[\mathbf{\Gamma}_{11}(0)]^{-1}\mathbf{x}_i = 1,$$

and,

$$\mathbf{x}_i^+[\mathbf{\Gamma}_{11}(1)]^{-1}\mathbf{x}_i = 0.$$

Those formulas are well known as the generalized eigenvalue problem. This system can be represented in a closed form through a generalized Rayleigh quotient:

$$\rho(\mathbf{x}_i) = \frac{\mathbf{x}_i^+[\mathbf{\Gamma}_{11}(1)]^{-1}\mathbf{x}_i}{\mathbf{x}_i^+[\mathbf{\Gamma}_{11}(0)]^{-1}\mathbf{x}_i},$$

i.e. the problem of the minimization of a quadratic form with a quadratic constraint. We remark that $[\mathbf{\Gamma}_{11}(1)]^{-1}$ is not, always, positive definite, then this equation have not a closed form solutions. Therefore, instead of $\mathbf{\Gamma}_{11}(1)$, a combination of this matrix with its transpose conjugate is used. However, we can use the conjugate gradient techniques to obtain the desired solutions. These systems are, still, available in the case of an additive white noise.

Using the same partition as above, in the next section, we estimate, directly, the signal parameters from a small part of the data.

5 BLIND MULTICHANNEL DECONVOLUTION

In this section, we propose an improved version of the technique presented in [3]. The idea of this approach is, by avoiding the noise effect and using, only, a part of observations, the estimation of the vector $\mathbf{z}$ such as,

$$\mathbf{z}^+\mathbf{x}_1(n) = \mathbf{z}^+\mathcal{X}\mathbf{s}(n) = e^{j\phi}\mathbf{s}(n-k),$$

where ϕ is a phase ambiguity inherent to the problem. we use a cost function to find this vector $\mathbf{z}$, in order to deconvolve the transmitted sequence. The vector $\mathbf{z}$ is estimate from the following quadratic form with a quadratic constraint:

$$\mathbf{z}^{+}[\mathcal{X}\mathcal{X}^{+}]\mathbf{z} = 1,$$

and,

$$\mathbf{z}^{+}[\mathcal{X}\mathcal{G}\mathcal{X}^{+}]\mathbf{z} = 0,$$

where $\mathcal{G}$ is a symetric matrix of rank $(N + M - 1)$ with a null subspace whose associated eigenvector is $x = [\mathbf{0}\ 1\ \mathbf{0}]$. In order to increase the algorithm's efficiency, an alternative implementation based on the following criterion is,

$$\min_{\mathbf{z}} \left(\frac{\mathbf{z}^{+}\mathcal{B}\mathcal{B}^{+}\mathbf{z}}{\mathbf{z}^{+}\mathcal{A}\mathbf{z}} \right),$$

where $\mathcal{B}\mathcal{B}^{+}$ is a positive definite matrix, with $\mathcal{B} = \mathcal{X}\mathcal{G}\mathcal{X}^{+}$ and $\mathcal{A} = \mathcal{X}\mathcal{X}^{+}$. The solution of this equation is the eigenvector corresponding to the minimum generalized eigenvalue of $(\mathcal{B}\mathcal{B}^{+}, \mathcal{A})$.

The main goal is to increase the efficiency of the original approach by using, solely, the $(N + M)$ elements of the observation vector. Also, the improved method allows a reduction of the noise and the complexity load.

6 PERFORMANCE EVALUATION

To study the performance of the presented approaches, we consider a stationary scenario. The channel coefficients are given in [1], the number of virtual channels is $L = 4$; the width of the temporal window is $N = 10$; the degree of the ISI is $M = 4$.

A Monte Carlo simulation was conducted to evaluate the performance of the presented methods. We varied the number of symbols (K) and the signal-to-noise ratio (SNR) in function of the normalized root mean-square error $(NRMSE)$ defined below. For each K and SNR, 100 Monte Carlo runs were conducted for different realisations. The number K of data samples used to estimate each $\mathbf{h}$ ranges from 100 to 1000. The normalized root mean-square error $(NRMSE)$ is employed as a performance measure of the input estimates:

$$NRMSE = \frac{1}{\| \mathbf{h} \|} \left(\frac{1}{T} \sum_{t=1}^{T} \| \widehat{\mathbf{h}}_t - \mathbf{h} \|^2 \right)^{1/2}, \text{ where } T \text{ is}$$

the number of trials (100 in our cases), $\widehat{\mathbf{h}}_t$ is the estimate of the inputs from the tth trial. The signal to noise ratio (SNR) is defined as $SNR = 10 \log_{10} \frac{E\{\|\mathcal{H}\mathbf{s}(k)\|^2\}}{E\{\|\mathbf{n}(k)\|^2\}}$.

These figures show that the channel identification by the linear operator or by the direct method have, slightly, the same performance as the well-known method, the subspace approach. In addtion, we have with the two first methods an important reduction of computational load because we, only, use a small part of data.

7 CONCLUSION

Several techniques were presented to identify the channel and/or to estimate the signal parameters. These methods are based on well-known methods of subspace and Tong's approach. These improvements allow a direct equalization and the identification of the channel using, only, a small part of data.

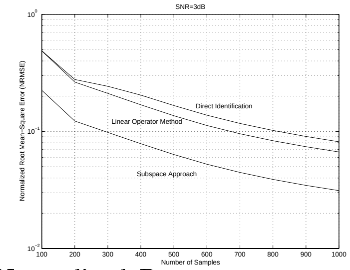

Figure 1: Normalized Root mean-square error of the channel parameters estimates in function of the number of samples.

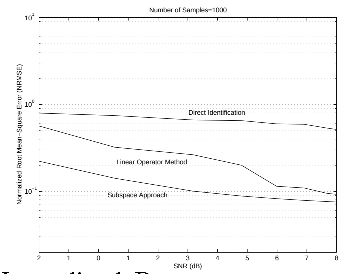

Figure 2: Normalized Root mean-square error of the channel parameters estimates in function of SNR.

References

[1] E. Moulines et al., "Subspace methods for the blind identification of multichannel FIR filters," IEEE Transactions on Signal Processing, vol. 43, no. 2, pp. 516-525, February 1995.

[2] L. Tong et al., "Blind identification and equalization based on second-order statistics: A time domain approach," IEEE Transactions on information Theory, vol. 40, no. 2, pp. 340-349, March 1994.

[3] S. Zazo, J. Borrallo, "A new pencil criterium for multichannel blind deconvolution in data communication systems", Proceedings of ICASSP'98, May 1998, pp. 2361-2364.

[4] M. Frikel, V. Barroso and J. Xavier, "Blind Channel Identification Using the Propagator Operator," Proceedings of II Conference in Telecommunication, April 14-16, 1999, Portugal.

[5] M. Frikel, V. Barroso and J. Xavier, "Blind Recursive Estimation of SIMO Channels", Proceedings IEEE SPAWC, Maryland, USA, May, 1999.

[6] K. Abed-Meraim et al., "Prediction Error Method for Second-Order Blind Identification," IEEE Transactions on Signal Processing, vol. 45, no. 3, pp. 694-705, March 1997.