2nd IEEE Workshop on Signal Processing Advances in Wireless Communications, May 9-12, 1999 - Annapolis, MD, USA, pp-275–278

A DETERMINISTIC SCHUR METHOD FOR MULTICHANNEL BLIND IDENTIFICATION

Luc Deneire and *Dirk T.M. Slock*

{deneire, slock}@eurecom.fr
Institut EURECOM, 2229 route des Crêtes, B.P. 193, F-06904 Sophia Antipolis Cedex, FRANCE

ABSTRACT

We address the problem of blind multichannel identification in a communication context. Using a deterministic model for the input symbols and only second order statistics, we develop a simple algorithm, based on the Generalized Schur algorithm to apply LDU decomposition of the covariance matrix of the received data. We show that this method leads to identification of the channel, up to a constant. Furthermore, the identification algorithm is shown to yield similar performance as the subspace method [7]. This paper complements [2] where we developed a stochastic Schur algorithm.

1. INTRODUCTION

Blind multichannel identification has received considerable interest over the last decade. In particular, second-order methods have raised a lot of attention, due to their ability to perform channel identification with relatively short data bursts. Among these methods, we can distinguish the deterministic methods, where the input symbols are considered deterministic and the stochastic methods, where the input symbols are considered stochastic. In a multiuser environment, using the deterministic model leads to a dynamical indeterminacy [3, 5] as opposed to the stochastic model which leads to the identification of the channel up to a unitary static mixture matrix [3]. Subsequent source separation can then be performed by other classical methods or by resorting to known symbols (i.e. performing semi-blind identification). On the other hand, use of deterministic methods lead to consistency in SNR, which can be an interesting feature.

We show that LDU decomposition of the covariance matrix leads to the identification of the channel (up to a scalar) and that performing this decomposition with a Schur algorithm yields good performance.

The work of Luc Deneire is supported by the EC by a Marie-Curie Fellowship (TMR program) under contract No ERBFMBICT950155 and the work of Dirk Slock by Laboratoires d'Electronique PHILIPS under contract LEP 95FAR008 and EURECOM Institute.

2. DATA MODEL AND NOTATIONS

Consider linear digital modulation over a linear channel with additive Gaussian noise. Assume that we have 1 transmitter at a certain carrier frequency and m antennas receiving mixtures of the signals. We shall assume that $m > 1$. The received signals can be written in the baseband as

$$y_i(t) = \sum_k a(k)h_i(t-kT) + v_i(t),\ i = 1,\cdots,m \quad (1)$$

where the $a(k)$ are the transmitted symbols, T is the common symbol period, $h_i(t)$ is the (overall) channel impulse response from the transmitter to receiver antenna i. Assuming the $\{a(k)\}$ and $\{v_i(t)\}$ to be jointly (wide-sense) stationary, the processes $\{y_i(t)\}$ are (wide-sense) cyclostationary with period T. If $\{y_i(t)\}$ is sampled with period T, the sampled process is (wide-sense) stationary.

We assume the channels to be FIR. In particular, after sampling we assume the (vector) impulse response from to be of length N. The discrete-time received signal can be represented in vector form as

$$\begin{aligned} \boldsymbol{y}(k) &= \sum_{i=0}^{N-1} \boldsymbol{h}(i)a(k-i) + \boldsymbol{v}(k) \\ &= \boldsymbol{H}\boldsymbol{A}_N(k) + \boldsymbol{v}(k); \boldsymbol{H} = [\boldsymbol{H}_1 \cdots \boldsymbol{H}_p] \end{aligned} \quad (2)$$

We consider additive temporally and spatially white Gaussian circular noise $\boldsymbol{v}(k)$ with $R_{vv}(k-i) = \mathrm{E}\{\boldsymbol{v}(k)\boldsymbol{v}^H(i)\} = \sigma_v^2 I_m \delta_{ki}$. Assume we receive M samples :

$$\boldsymbol{Y}_M(k) = \mathcal{T}_M(\boldsymbol{H})\ \boldsymbol{A}_{N+p(M-1)}(k) + \boldsymbol{V}_M(k) \quad (3)$$

where $\boldsymbol{Y}_M(k) = \left[\boldsymbol{Y}^H(k)\cdots\boldsymbol{Y}^H(k-M+1)\right]^H$ and $\boldsymbol{V}_M(k)$ is defined similarly whereas $\mathcal{T}_M(\boldsymbol{H})$ is the multichannel convolution matrix of $\boldsymbol{H}$, with M block lines and $\mathcal{T}_M(\boldsymbol{H})$ is block Toeplitz. The input symbols are deterministic.

3. LDU FACTORIZATION OF A COVARIANCE MATRIX

From $R_{YY} = \mathrm{E}\{YY^H\}$, under the identifiability conditions, we can identify σ_v^2 as the singular vector corresponding to the minimum singular value of R_{YY}. Let $\widetilde{Y}$ be the prediction error, then, as Y can be perfectly predicted in the absence of noise, the covariance of the error can be written as

$$R_{YY} - R_{Y\tilde{Y}}R^{\#}_{\tilde{Y}\tilde{Y}}R_{\tilde{Y}Y} = 0 \Rightarrow R_{Y\tilde{Y}}R^{\#}_{\tilde{Y}\tilde{Y}}R_{\tilde{Y}Y} = U^H DU \tag{4}$$

where $^{\#}$ denotes a generalized inverse.

Consider we perform a block triangularization, examination of the rank of $R_{YY} - \sigma_v^2 I = \mathcal{T}(\boldsymbol{H})\mathcal{T}^H(\boldsymbol{H})\sigma_a^2$ leads to, for block i

$$\text{rank}\,(D_i) = \begin{cases} = 1 & , i \geq \underline{L} \\ = m - \underline{m} \in \{2, \ldots, m\} & , i = \underline{L}-1 \\ = m & , i < \underline{L}-1 \end{cases} \tag{5}$$

where $\underline{m} = (m-1)\underline{L} - N - 1 \in \{0, 1, \ldots, m\}$ and $\underline{L} = \left\lceil \frac{N-1}{m-1} \right\rceil$.

As the prediction of $\boldsymbol{Y} = \boldsymbol{Y}_M(k)$ is perfect from instant $\underline{L}+1$ on, in $t > \underline{L}$, $\tilde{\boldsymbol{y}}(k-t)$ contains the emitted symbols, apart from a unitary matrix, which is consistent with the rank profile of D_i. Furthermore, denoting $U(i,j)$ as the (i,j) block of U, $U^H = R_{Y\tilde{Y}}$ implies:

$$\begin{aligned} U^H(i,j) &= \mathrm{E}\left\{\boldsymbol{y}(k-i)\tilde{\boldsymbol{y}}^H(k-j)\right\} \\ &= \sum_{l=0}^{N-1} \boldsymbol{H}(l)\mathrm{E}\left\{\boldsymbol{a}(k-i-l)T\, \boldsymbol{a}^H(k-j)\right\} \\ &\quad \text{for } i,j > \underline{L} \\ &= \alpha\, \boldsymbol{H}(j-i)\sigma_a^2 \end{aligned} \tag{6}$$

where α is a constant scalar. Hence, we can identify the channel, up to a constant, by triangularization of the covariance matrix of the received signal.

4. APPLYING THE GENERALIZED SCHUR ALGORITHM TO LDU FACTORIZATION OF $\widehat{R}_{YY}$

In [2], we used the so-called "biased estimator" of the correlation sequence

$$\widehat{R}_{YY}(i) = \frac{1}{M}\sum_{t=1}^{M} \boldsymbol{y}(t)\boldsymbol{y}^H(t+i),$$

because the matrix formed with these estimators is, by construction, block Toeplitz and definite positive. Use of this estimator leads to a stochastic method, which has the advantage of being robust to the overestimation of the order and, in a multiuser environment, to give the channel up to an unitary constant matrix.

In this paper, we use the "sample covariance matrix" ($\widehat{R}_{YY} = \frac{1}{M}\sum_{t=0}^{M-1}\boldsymbol{Y}(t)\boldsymbol{Y}^H(t)$, where $\boldsymbol{y}(k) = 0$ for $k => M$), which has a displacement rank of $2(m+1)$, having thus a different structure as the true covariance matrix.

5. USE OF THE GENERALIZED SCHUR ALGORITHM

5.1. Some basics

The displacement of a $n \times n$ Hermitian matrix is defined as $\nabla R \triangleq R - Z_\nu R Z_\nu^H$, where Z_ν is a $n \times n$ lower shift matrix with ones on the ν^{th} sub-diagonal [1]. The rank r of ∇R is called the displacement rank and can be shown to be equal to $2m$ for the covariance matrix $R_{YY}(-\sigma_v^2 I)$. Moreover, we can factor ∇R_{YY} as $\nabla R_{YY} = G\Sigma G^H$ where $\Sigma = (I_m \oplus -I_m)$ is called the signature matrix and G the generator of R_{YY}. One can easily check that, denoting the blocs $r_i \triangleq R_{YY}^{-1/2}(0)R_{YY}(i)$, R_{YY} of size $Km \times Km$ and $\mathcal{R}_{YY} = [r_{i-j}]_{i,j}$.

$$\nabla\mathcal{R}_{YY} = \begin{bmatrix} r_0 & 0 \\ r_1 & r_1 \\ \vdots & \vdots \\ r_{K-1} & r_{K-1} \end{bmatrix} \begin{bmatrix} I_m & 0 \\ 0 & -I_m \end{bmatrix} \begin{bmatrix} r_0 & 0 \\ r_1 & r_1 \\ \vdots & \vdots \\ r_{K-1} & r_{K-1} \end{bmatrix}^H$$

Proceeding by block, the generalized Schur algorithms starts with the generator $G^{(0)} = G$, forms

$$\begin{aligned} G^{(1)^H} &= S^{(1)} \begin{bmatrix} 0 & r_0^{(0)} & r_1^{(0)} & \cdots & r_{K-1}^{(0)} \\ 0 & r_1^{(0)} & r_2^{(0)} & \cdots & r_K^{(0)} \end{bmatrix} \\ &= \begin{bmatrix} 0 & r_0^{(1)} & r_1^{(1)} & \cdots & r_{K-1}^{(1)} \\ 0 & 0 & \tilde{r}_2^{(1)} & \cdots & \tilde{r}_K^{(1)} \end{bmatrix} \end{aligned} \tag{7}$$

where $S^{(1)}$ is a block hyperbolic Householder transformation (such that it is Σ unitary: i.e. $S^{(1)}\Sigma S^{(1)^H} = \Sigma$). Then $G^{(1)}$ is the generator of the Schur complement of $\mathcal{R}_{YY}$ with respect to r_0. Continuing this process further, we get

$$\mathcal{R}_{YY} = U^H DU \text{ where } U = \begin{bmatrix} r_0^{(0)} & r_1^{(0)} & \cdots & r_{K-1}^{(0)} \\ 0 & r_0^{(1)} & \ddots & r_{K-2}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_0^{(K-1)} \end{bmatrix}$$

[1] broader definitions can be found in [6, 1].

5.2. Applying it to LDU factorization of $\widehat{R}_{YY}$

We can write $\widehat{R}_{YY} = \mathcal{Y}\mathcal{Y}^H$ where

$$\mathcal{Y} = \begin{bmatrix} y(0) & y(1) & \cdots & y(M-K+1) \\ y(1) & y(2) & \cdots & y(M-K+2) \\ \vdots & \vdots & \ddots & \vdots \\ y(K-1) & y(K) & \cdots & y(M) \end{bmatrix}$$

and $\mathcal{Y}$ can be written as $\mathcal{T}(\boldsymbol{H})\mathcal{A}$ for some Toeplitz matrix $\mathcal{A}$. Hence, $\widehat{R}_{YY} = \mathcal{T}(\boldsymbol{H})\mathcal{A}\mathcal{A}^H\mathcal{T}(\boldsymbol{H})^H$ which implies that $\widehat{R}_{YY}$ has a displacement rank $2m+2$ and it's generators are :

$$\begin{bmatrix} r_0 & y(M-K+1) & 0 & 0 \\ r_1 & y(M-K+2) & r_1 & y(0) \\ \vdots & \vdots & \vdots & \vdots \\ r_{K-1} & y(M) & r_{K-1} & y(K-2) \end{bmatrix}$$

Partitioning the received signal matrix as

$$\left[\frac{\mathcal{Y}_1}{\mathcal{Y}_2}\right] = \left[\begin{array}{c|c} \mathcal{T}_1 & 0 \\ \hline \mathcal{T}_x & \mathcal{T}_2 \end{array}\right]\left[\frac{\mathcal{A}_1}{\mathcal{A}_2}\right] \quad (8)$$

we can express the estimation error on $\mathcal{Y}_2$ as $\widetilde{\mathcal{Y}}_2 = \mathcal{Y}_2 - \mathcal{Y}_2\mathcal{Y}_1^H\left(\mathcal{Y}_1\mathcal{Y}_1^H\right)^{-1}\mathcal{Y}_1 = \mathcal{Y}_2\mathrm{P}^{\perp}_{\mathcal{Y}_1^H}$, where $\mathrm{P}^{\perp}_{\mathcal{Y}_1^H}$ is the orthogonal projection on the span of $\mathcal{Y}_1$. From 8, $\mathrm{P}^{\perp}_{\mathcal{Y}_1^H} = \mathrm{P}^{\perp}_{\mathcal{A}_1^H}$, hence

$$\widetilde{\mathcal{Y}}_2\widetilde{\mathcal{Y}}_2^H = \mathcal{T}_2\mathcal{A}_2\mathrm{P}^{\perp}_{\mathcal{A}_1^H}\mathcal{A}_2^H\mathcal{T}_2^H \quad (9)$$

where $\widetilde{\mathcal{Y}}_2\widetilde{\mathcal{Y}}_2^H$ is the Schur complement of the lower block in $M\widehat{R}_{YY}$.

The Toeplitz nature of $\mathcal{A}$ shows that the displacement rank of $\mathcal{A}_2\mathrm{P}^{\perp}_{\mathcal{A}_1^H}\mathcal{A}_2^H$ is 4, hence the displacement rank of $\widetilde{\mathcal{Y}}_2\widetilde{\mathcal{Y}}_2^H$ is also 4, provided the size of $\mathcal{T}_1$ is bigger than $\underline{L} \times \underline{L}$ blocks.

From this structure, one can write the Schur complement in the displacement form as

$$\widetilde{\mathcal{Y}}_2\widetilde{\mathcal{Y}}_2^H = \sum_{i=1}^{4}(\mathcal{T}_2L_i)(\mathcal{T}_2L_i)^H$$

and the generators can be seen as the convolution of the channel with the prediction filter of the data. Hence, once we have found the generators, it suffices to take the generator corresponding to the maximum positive eigenvalue of the generator matrix and separate the channel part from the prediction part, by any suitable mean (e.g. subspace fitting or less complex methods). This completes the deterministic Schur algorithm we implemented.

Remarks

This displacement structure could lead to more involved algorithms, using the 4 generators to find the channel, leading to marginally better performance. From the structure here above, it can be clearly seen that the extension to the multiuser case is "trivial", as the displacement rank of $\mathcal{A}_2\mathrm{P}^{\perp}_{\mathcal{A}_1^H}\mathcal{A}_2^H$ is then $2p+2$, where p is the number of users. But the very deterministic nature of the method leads then to identifiability up to a dynamical matrix (see [3]). Another way of performing a stochastic Schur channel identification is then to assume that $\mathcal{A}_2\mathrm{P}^{\perp}_{\mathcal{A}_1^H}\mathcal{A}_2^H = I$, which leads to $L_i = I\delta(i)$.

SCHUR ESTIMATION PROCEDURE

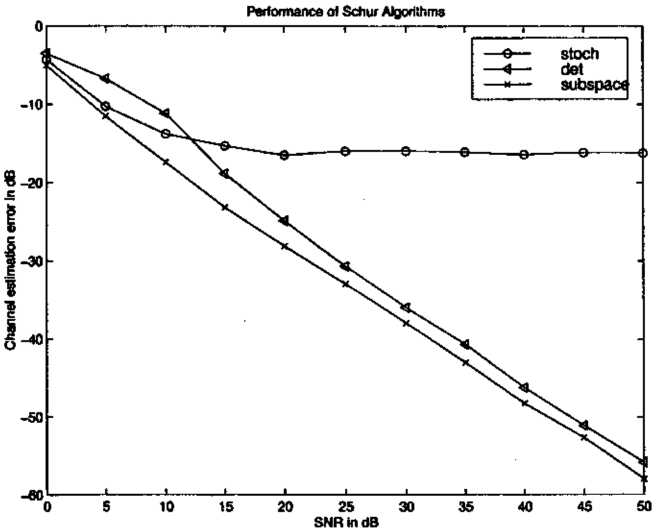
1. Calculate the sample covariance matrix $\widehat{R}_{YY}$ then $\widehat{R}_{YY} - \lambda_{min}(\widehat{R}_{YY})I$
2. Calculate $\widehat{R}_{YY}^{1/2}$.
3. Proceed with the Schur iterations as in (7) until the $\underline{L}^{\text{th}}$ iteration.
4. Either
   - collect the first columns of the generator.
   - calculate the g, the eigenvector of the generator corresponding to the maximum eigenvalue.
   - calculate the channel by any deterministic method from g.

6. SIMULATIONS

In order to evaluate the performance of the algorithms, we have computed the Normalized MSE (NMSE) on the estimated channels, averaged over 100 Monte Carlo runs. We have used a randomly generated channel with channel length $N = 5$, and $m = 3$ sub-channels. The symbols are i.i.d. BPSK and the data length is $M = 100$. The constant scalar has been estimated afterwards as $\alpha = \frac{\boldsymbol{H}_t\boldsymbol{H}_t^H}{\widehat{\boldsymbol{H}}_t\boldsymbol{H}_t^H}$ where $\boldsymbol{H}_t = [\boldsymbol{h}(0)^T \ldots \boldsymbol{h}(N-1)^T]$

We evaluate the performance of the deterministic and stochastic Schur algorithms and of the subspace algorithm [7].

For the stochastic Schur algorithm, we used the procedure detailed in [2]. For the deterministic algorithm, we explicitly calculated the displacement matrix of the Schur complement in the sample covariance matrix, faster algorithms can be found in [8, 4]. Further on, we computed the maximum eigenvector of this matrix and performed a subspace fitting on this vector to get the actual channel.



Comparison between Schur and subspace algorithms

Curves show that the deterministic Schur algorithm gives results comparable to the subspace algorithm, while the stochastic Schur algorithm suffers from the well known flooring effect at high SNR.

7. CONCLUSIONS

We have introduced a deterministic Schur method to identify multichannels blindly. Using the generalized Schur algorithm to find the generators of a Schur complement of the sample covariance matrix, we find these generators to be the convolution between the channel and the source prediction filters. Hence, we can deduce the channel exactly in the noiseless case.

This algorithm is recursive in order, when using the generalized Schur algorithm adapted for singular matrices, and can be coupled to a channel length estimator (and source detector in the multiuser case) by examining the diagonal of the LDU of $R_{YY} - \sigma_v^2 I$. The performance of this algorithm is shown to be close to the subspace method, which is near optimal [7].

Extension to the multiuser case can easily be done, but lead to indeterminacy problems [3] (i.e. there remains a convolutive mixture). Hence, further work should lead to introduce some mix between the deterministic algorithm and the stochastic one, using a softer condition than $A_2 \mathrm{P}^{\perp}_{A_1^H} A_2^H = I$, in the form on a soft condition on the generators of this matrix.

8. REFERENCES

[1] P. COMON. Structured matrices and inverses. In A. Bojanczyk and G. Cybenko, editors, *Linear Algebra for Signal Processing*, volume 69 of *IMA Volumes in Mathematics and its Applications*, pages 1–16. Springer Verlag, 1995.

[2] Luc Deneire and Dirk Slock. "A Schur method for multiuser multichannel blind identification,". In *submitted to ICASSP99*.

[3] Luc Deneire and Dirk Slock. "Identifiability conditions for blind and semi-blind multiuser multichannel identification". In *9th IEEE Signal Processing Workshop On Statistical Signal And Array Processing*, Portland, Oregon, USA, September 1998.

[4] K.A Gallivan, S. Thirumalai, P. Van Dooren, and V. Vermaut. "High Performance Algorithms for Toeplitz and block Toeplitz matrices". *Linear Algebra and its Applications*, 1994.

[5] A. Gorokhov and Ph. Loubaton. "Subspace based techniques for second order blind separation of convolutive mixtures with temporally correlated sources". *IEEE Trans. on Circuits and Systems*, July 1997.

[6] J. Chun. *"Fast Array Algorithms for Structured Matrices"*. PhD thesis, Information Systems Laboratory, Stanford University, June 1989.

[7] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue. "Subspace Methods for the Blind Identification of Multichannel FIR filters". *IEEE Trans. on Signal Processing*, 43(2):516–525, February 1995.

[8] Ali H. Sayed and Thomas Kailath. "A Look-ahead Block Schur Algorithm for Toeplitz-Like Matrices". *SIAM J. Matrix Anal. Appl.*, 16(2), April 1995.