A STOCHASTIC SUBSPACE ALGORITHM FOR BLIND CHANNEL IDENTIFICATION IN NOISE FIELDS WITH UNKNOWN SPATIAL COLOR

Piet Vandaele, Marc Moonen

ESAT - Katholieke Universiteit Leuven
K. Mercierlaan 94, 3001 Heverlee -Belgium
tel 32/16/32 17 99 fax 32/16/32 19 70
piet.vandaele@esat.kuleuven.ac.be marc.moonen@esat.kuleuven.ac.be

ABSTRACT

In this paper, the blind channel identification problem is formulated in a stochastic state space framework. Starting from a state space model we present a preprocessing step based on two orthogonal subspace projections. Using these orthogonal projections, we derive an algorithm for blind channel estimation which is insensitive to the spatial color of the noise. The performance of this new algorithm is demonstrated through simulation examples.

1. INTRODUCTION

Blind channel identification based on second order statistics or equivalent deterministic properties has been an active area of research during the last years [2]. Most algorithms developed up till now assume that the additive noise is spatially white or that the spatial covariance of the noise is known. An exception is [1], in which the original subspace algorithm [3] is modified to cope with noise of unknown spatial covariance. For a proper operation of the algorithm [1] it is required that the number of observation channels exceeds two and that the channel order exceeds one.
Here we present a new stochastic subspace algorithm for blind channel identification. The algorithm is based on the concept of orthogonal projections and is closely related to the theory of [4]. It has the same restrictions for the number of channels and the channel order as the algorithm [1]. However, simulation results show that the new algorithm has a better performance than the algorithm of [1]. The main contribution of this paper lies in the application of the stochastic subspace ideas of [4] to the blind identification problem.

Piet Vandaele is a Research Assistant supported by the Flemish I.W.T. (the Flemish Institute for the Advancement of Scientific-Technological Research in Industry). Marc Moonen is a Research Associate with the Fund for Scientific Research - Flanders (Belgium) (F.W.O.). This research work was carried out at the ESAT laboratory of the Katholieke Universiteit Leuven, in the framework of a Concerted Action Project of the Flemish Community, entitled *Model-based Information Processing Systems* (GOA/MIPS/95/99/3) as well as the IT-program of the I.W.T., *Integrating Signal Processing Systems* (ITA/GBO/T23) and was partially sponsored by IMEC (Flemish Interuniversity Microelectronics Center) and IUAP P4-02 (1997-2001): Modeling, Identification, Simulation and Control of Complex Systems. The scientific responsibility is assumed by its authors.

2. DATA MODEL

Consider the following single input- M output channel model:

$$\mathbf{y_k} = \begin{bmatrix} \mathbf{h_L} & \dots & \mathbf{h_0} \end{bmatrix} \begin{bmatrix} x[k-L] \\ \vdots \\ x[k] \end{bmatrix} + \mathbf{n_k} \qquad (1)$$

Here $\mathbf{y_k}$ represents the channel output vector at time k, i.e. the outputs of M receiving antennas. The channel input is denoted as $x[k]$ and the vector FIR channel as $\mathbf{h_k}$. *The noise vector* $\mathbf{n_k}$ *is assumed to be temporally white but can be spatially colored.*
First we rewrite the data model of equation 1 into a forward and backward stochastic state space model.

2.1. Forward state space model

Define the state vector $\mathbf{x_k} = \begin{bmatrix} x[k-L] & \cdots & x[k-1] \end{bmatrix}$ then we obtain the following state and output equation

$$\begin{aligned} \mathbf{x_{k+1}} &= A \cdot \mathbf{x_k} + \underbrace{\mathbf{e_L} \cdot x[k]}_{\mathbf{w_k}} \\ \mathbf{y_k} &= \underbrace{\begin{bmatrix} \mathbf{h_L} & \cdots & \mathbf{h_1} \end{bmatrix}}_{C} \cdot \mathbf{x_k} + \underbrace{(\mathbf{h_0} \cdot x[k] + \mathbf{n_k})}_{\mathbf{v_k}} \end{aligned}$$

where A is a matrix which has unit entries on its first super diagonal and zero entries everywhere else, and $\mathbf{e_L}$ is the last column of the $L \times L$ unity matrix. The (unknown) input signal is treated as a random binary white noise sequence. The conditions under which this data model is minimal (i.e. both observable and controllable) will be detailed in the next section.

2.2. Backward state space model

Define the state vector $\mathbf{z_k} = \begin{bmatrix} x[k-L+1] & \cdots & x[k] \end{bmatrix}$ then we obtain the following state and output equation

$$\begin{aligned} \mathbf{z_{k-1}} &= A^H \cdot \mathbf{z_k} + \underbrace{\mathbf{e_1} \cdot x[k-L]}_{\mathbf{w_k^b}} \\ \mathbf{y_k} &= \underbrace{\begin{bmatrix} \mathbf{h_{L-1}} & \cdots & \mathbf{h_0} \end{bmatrix}}_{G^H} \cdot \mathbf{z_k} + \underbrace{(\mathbf{h_L} \cdot x[k-L] + \mathbf{n_k})}_{\mathbf{v_k^b}} \end{aligned}$$

with $\mathbf{e_1}$ the first column of a $L \times L$ unity matrix. Note that $\mathbf{z_k}$ equals $\mathbf{x_{k+1}}$.

3. ORTHOGONAL PROJECTIONS

We construct the following two block Hankel matrices, called 'past outputs' and 'future outputs' and respectively denoted as Y_p and Y_f:

$$\begin{aligned} Y_p &= Y_{1|i} = \begin{bmatrix} \mathbf{y_1} & \cdots & \mathbf{y_j} \\ \vdots & & \vdots \\ \mathbf{y_i} & \cdots & \mathbf{y_{j+i-1}} \end{bmatrix} \\ Y_f &= Y_{i+1|2\cdot i} = \begin{bmatrix} \mathbf{y_{i+1}} & \cdots & \mathbf{y_{j+i}} \\ \vdots & & \vdots \\ \mathbf{y_{2\cdot i}} & \cdots & \mathbf{y_{j+2\cdot i-1}} \end{bmatrix} \end{aligned}$$

The subscripts refer to the time indices in the first column (the number of columns is always fixed to j). Further define the $M \cdot i \times (L+i)$ matrix $\mathcal{H}_i$:

$$\mathcal{H}_i = \begin{bmatrix} \mathbf{h_L} & \cdots & \mathbf{h_0} & & \\ & \ddots & & \ddots & \\ & & \mathbf{h_L} & \cdots & \mathbf{h_0} \end{bmatrix}$$

and $X_{1-L|i}$ as:

$$X_{1-L|i} = \begin{bmatrix} x[1-L] & \cdots & x[j-L] \\ \vdots & & \vdots \\ x[i] & \cdots & x[i+j-1] \end{bmatrix}$$

then,

$$\begin{aligned} Y_p &= \mathcal{H}_i \cdot X_{1-L|i} + N_p \\ Y_f &= \mathcal{H}_i \cdot X_{i+1-L|2\cdot i} + N_f \end{aligned}$$

with $X_{i+1-L|2\cdot i}$ defined similarly as $X_{1-L|i}$ and N_p,N_f defined in a similar way as Y_p and Y_f. For the derivations to follow we make the following assumptions:

H1 $\mathcal{H}_i$ has full column rank for $i \geq L$.

H2 $j \to \infty$. This allows to apply stochastic properties of the data model.

H3 Source symbols are uncorrelated, i.e. $E\{x[k] \cdot x[l]\} = \delta_{kl}$ and noise and symbols are uncorrelated.

H4 The forward and backward state space models are both observable and controllable.

We first show that assumption H1 automatically implies H4. For the forward state space model we define the extended observability matrix Γ_i:

$$\begin{aligned} \Gamma_i &= \begin{bmatrix} C & C\cdot A & \cdots & C\cdot A^{i-1} \end{bmatrix}^T \\ &= \mathcal{H}_i(:,1:L) \end{aligned} \quad (2)$$

The pair $\{A, C\}$ is observable if Γ_i has full column rank, which is clearly the case under assumption H1. Straightforward calculation also yields that the pair $\{A, Q^{1/2}\}$ is controllable (i.e. that all dynamical modes of the system are excited by the (unknown) input signal):

$$\text{rank}(\begin{bmatrix} A^{i-1} \cdot Q^{1/2} & \cdots & A \cdot Q^{1/2} & Q^{1/2} \end{bmatrix}) = L$$

with Q:

$$Q = E\{\mathbf{w_k} \cdot \mathbf{w_k}^H\} = \begin{bmatrix} 0_{L\times(L-1)} & \mathbf{e_1} \end{bmatrix}$$

Similar expressions can be derived for the backward model. We now define a number of orthogonal projections. Similar to [4], we orthogonally project the future outputs Y_f onto the past outputs Y_p, then under the assumptions H2 to H4:

$$\begin{aligned} \mathcal{O}_i &= (Y_f/Y_p) = Y_f \cdot Y_p^H \cdot (Y_p \cdot Y_p^H)^\dagger \cdot Y_p \\ &= \Gamma_i \cdot \hat{X}_{i+1-L|i} \end{aligned} \quad (3)$$

This equation is valid whatever the spatial color of the noise, and represents a rank L model as will now be explained. Γ_i is a block upper triangular, block Toeplitz matrix as defined in equation 2:

$$\Gamma_i = \mathcal{H}_i(:,1:L) = \begin{bmatrix} \mathbf{h_L} & \cdots & \mathbf{h_1} \\ & \ddots & \vdots \\ & & \mathbf{h_L} \\ 0 & \cdots & 0 \end{bmatrix}$$

$\hat{X}_{i+1-L|i}$ contains the non-steady state Kalman filter state estimates of the exact state sequence $\begin{bmatrix} \mathbf{x_{i+1}} & \cdots & \mathbf{x_{i+j}} \end{bmatrix}$ of the forward stochastic model [4], i.e.

$$\hat{X}_{i+1-L|i} = \begin{bmatrix} \mathbf{\hat{x}_{i+1}} & \cdots & \mathbf{\hat{x}_{i+j}} \end{bmatrix} \quad (4)$$

Now we make a number of important remarks concerning $\hat{X}_{i+1-L|i}$:

- The matrix $\hat{X}_{i+1-L|i}$ will be approximately Hankel. The Hankel structure will be better approximated whenever the noise level decreases.
- Computing $\hat{X}_{i+1-L|i}$ from an orthogonal projection can be viewed as generating a state sequence by a bank of non-steady state Kalman filters working in parallel on each of the columns of the block Hankel matrix of past outputs Y_p,

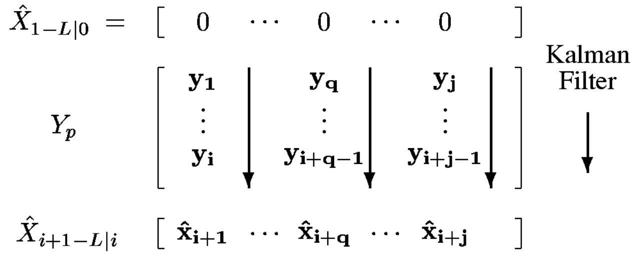

Figure 1: *Kalman filter state estimates based upon i measurements of $\mathbf{y_k}$. If the system matrices were known, the state $\hat{\mathbf{x}}_{\mathbf{i+q}}$ could be determined from a Kalman filter as follows: Start the filter at time q, with an initial state estimate 0. Now iterate the Kalman filter over i time steps (the vertical arrow down). The Kalman filter will then return a state estimate $\hat{\mathbf{x}}_{\mathbf{i+q}}$. This procedure could be repeated for each of the j columns, and thus we speak of a* bank *of Kalman filters. The observation here is that the system matrices do not have to be known to determine the state sequence $\hat{X}_{i+1-L|i}$. It can be determined directly from the output data, see [4].*

see [4] and figure 1. The bank of Kalman filters runs in a vertical direction (over the columns). They thus only use partial input-output information: i.e. the Kalman filter generating the estimate of $\hat{\mathbf{x}}_{\mathbf{i+q}}$ will only use i output measurements $\mathbf{y_q}, \dots, \mathbf{y_{i+q-1}}$ instead of all the output measurements $\mathbf{y_1}, \dots, \mathbf{y_{i+q-1}}$. $\hat{\mathbf{x}}_{\mathbf{i+q}}$ is thus the optimal one-step-ahead predicted state given the measurements of the outputs $\mathbf{y_q}, \dots, \mathbf{y_{i+q-1}}$.
Similarly we project the past outputs onto the future outputs:

$$\begin{aligned} \mathcal{B}_i &= (Y_p/Y_f) = Y_p \cdot Y_f^H \cdot (Y_f \cdot Y_f^H)^\dagger \cdot Y_f \\ &= \Delta_i^H \cdot \hat{\hat{X}}_{i+1-L|i} \end{aligned} \quad (5)$$

In this last equation Δ_i is the reversed extended stochastic controllability matrix:

$$\begin{aligned} \Delta_i^H &= \begin{bmatrix} A^{i-1} \cdot G & A^{i-2} \cdot G & \cdots & G \end{bmatrix}^H \\ &= \mathcal{H}_i(:, i+1 : L+i) = \begin{bmatrix} 0 & \cdots & 0 \\ \mathbf{h_0} & & \\ \vdots & \ddots & \\ \mathbf{h_{L-1}} & \cdots & \mathbf{h_0} \end{bmatrix} \end{aligned}$$

$\mathcal{H}_i(:, i+1 : L+i)$ is block lower triangular, block Toeplitz and contains the last L columns of $\mathcal{H}_i$. $\hat{\hat{X}}_{i+1-L|i}$ equals

$$\begin{aligned} \hat{\hat{X}}_{i+1-L|i} &= \begin{bmatrix} \hat{\mathbf{z}}_{\mathbf{i}} & \cdots & \hat{\mathbf{z}}_{\mathbf{i+j-1}} \end{bmatrix} \\ &= \begin{bmatrix} \hat{\hat{\mathbf{x}}}_{\mathbf{i+1}} & \cdots & \hat{\hat{\mathbf{x}}}_{\mathbf{i+j}} \end{bmatrix} \end{aligned} \quad (6)$$

Here $\hat{\mathbf{z}}_{\mathbf{i+q}}$ is the optimal one-step-backward predicted state given the measured outputs $\mathbf{y_{i+q+1}}, \dots, \mathbf{y_{2 \cdot i+q}}$. The matrix $\hat{\hat{X}}_{i+1-L|i}$ will only be approximately Hankel. The matrices $\hat{\hat{X}}_{i+1-L|i}$ and $\hat{X}_{i+1-L|i}$ form estimates of the same row space $X_{i+1-L|i}$. However, the two estimates will only coincide (and be exact) in the noiseless case.
An interesting feature of orthogonal projections is that they can be computed efficiently using a LQ decomposition, see [4].
Note that $\mathcal{B}_i$ and $\mathcal{O}_i$ will contain exactly $M \cdot L$ nonzero rows, this thus provides a way to estimate the channel order L. From here on we drop the zero rows and redefine $\mathcal{O}_i$ and $\mathcal{B}_i$ as:

$$\begin{aligned} \mathcal{O}_i &\leftarrow \mathcal{O}_i(1 : L \cdot M, :) \\ \mathcal{B}_i &\leftarrow \mathcal{B}_i((i-L) \cdot M + 1 : i \cdot M, :) \end{aligned}$$

From now on we will thus assume that $i = L$.
In the next section we present an algorithm that estimates the channel coefficients from both orthogonal projections. Since neither of the two projection equations contain all channel parameters, the two projections will have to be combined.

4. ALGORITHM

First we compute the matrices $\mathcal{O}_i$ and $\mathcal{B}_{i+L}$ by performing the appropriate orthogonal projections and add the results:

$$\mathcal{O}_i + \mathcal{B}_{i+L} = Y_{i+1|2 \cdot i}/Y_{1|i} + Y_{L+1|L+i}/Y_{i+L+1|2 \cdot i+L}$$
$$= \underbrace{\begin{bmatrix} \mathcal{H}_L(:, 1 : L) & \mathcal{H}_L(:, L+1 : 2 \cdot L) \end{bmatrix}}_{\mathcal{H}_L} \cdot \begin{bmatrix} \hat{X}_{i+1-L|i} \\ \hat{\hat{X}}_{i+1|i+L} \end{bmatrix}$$

Since the matrix $\begin{bmatrix} \hat{X}_{i+1-L|i} & \hat{\hat{X}}_{i+1|i+L} \end{bmatrix}^T$ has full row rank, the matrices $\mathcal{O}_i + \mathcal{B}_{i+L}$ and $\mathcal{H}_L$ share the same column space. Theorem 2 of [1], which is repeated below, then points out under which conditions, the channel parameters can be uniquely determined from this column space.

Theorem 1 *If $M \geq 3$, $L \geq 2$ and the assumptions H1-H3 hold, then the equation*

$$U_n^H \cdot \mathcal{H}_L = 0 \quad (7)$$

uniquely determines the channel $[\mathbf{h_0} \dots \mathbf{h_L}]$ (up to a scaling factor). U_n is the left null space of $\mathcal{H}_L$.

The left null space of $\mathcal{H}_L$, can be computed from the SVD of $\mathcal{O}_i + \mathcal{B}_{i+L}$. The channel parameters can then be retrieved by rewriting equation 7 as a function of the channel parameters and solving the problem using e.g. a quadratic non-triviality constraint, as in [3].

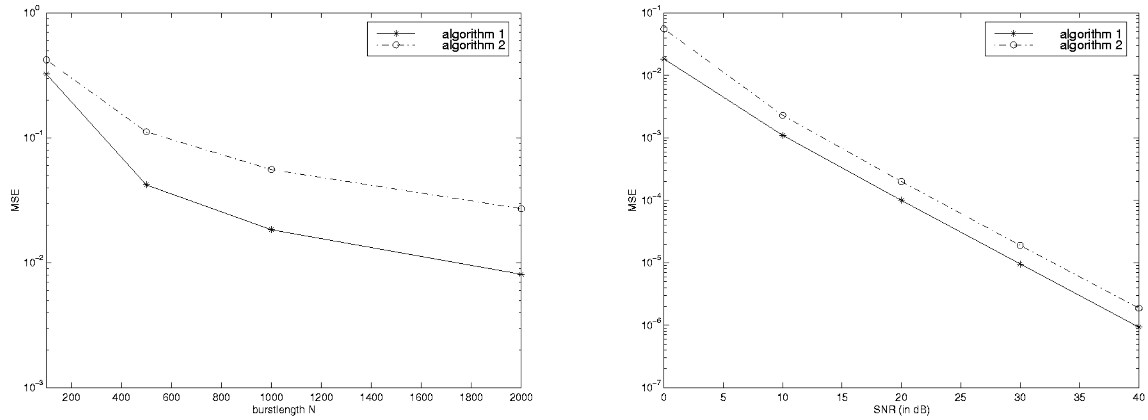

Figure 2: *MSE for varying burst length, SNR = 0 dB (left), MSE for varying SNR,* $N = 1000$ *(right)*

5. SIMULATION RESULTS

We simulate with a randomly generated (complex) channel of order 3:

$$\left[\begin{array}{ccc} \mathbf{h_L} & \ldots & \mathbf{h_0} \end{array}\right] =$$

$$\left[\begin{array}{cccc} 0.11+1.11i & -1.21-0.61i & 0.07-0.83i & -1.26+0.77i \\ 0.55+0.38i & 0.77-0.38i & -0.58+0.41i & -0.19+0.19i \\ 0.58+1.72i & -0.18-1.27i & 0.77+1.21i & -1.01+0.33i \end{array}\right]$$

so $M = 3$, $L = 3$. The modulation format is BPSK and the SNR is defined as:

$$SNR = 10 \cdot \log_{10} \frac{E\{\| \sum_{r=0}^{L} \mathbf{h_r} \cdot x[k-r] \|^2\}}{E\{\|\mathbf{n_k}\|^2\}}$$

The noise covariance matrix R_n is defined as:

$$R_n = \sigma^2 \cdot B \cdot B^H \text{ with } B = \left[\begin{array}{ccc} 1 & 0.7 & 0.7^2 \\ 0.7 & 1 & 0.7 \\ 0.7^2 & 0.7 & 1 \end{array}\right]$$

As a performance measure we use the mean square error of the channel estimate:

$$MSE = E\{\frac{\|H - \hat{H}\|_F^2}{\|H\|_F^2}\}$$

where $\| \cdot \|_F$ denotes the Frobenius norm. In all simulations we assume that the channel order is known and choose $i = L$. We compare the performance of our algorithm with algorithm [1], from now on referred to as respectively algorithm 1 and 2.

In a first simulation we investigate the influence of the burst length N on the performance of the two algorithms. The SNR was fixed to 0 dB. The lhs of figure 2 shows the results. For the new algorithm 1, the MSE decreases more rapidly with increasing burst length than for the original algorithm 2. In all situations algorithm 1 outperforms algorithm 2.

Next we vary the SNR and keep the burst length fixed at $N = 1000$. The rhs of figure 2 shows the results. Except for the lowest SNR, the noise level influences the MSE of both algorithms in an equal way.

6. CONCLUSIONS

In this paper we have presented a new blind equalization algorithm based on a stochastic state space description of the data model and orthogonal subspace projections. The algorithm is robust to the spatial color of the noise. The algorithm is similar to the algorithm of [1] but has better performance. Other algorithms combining the two orthogonal projections in a different way are the subject of current research.

7. REFERENCES

[1] K. Abed-Meraim, Y. Hua, P. Loubaton, and E. Moulines. Subspace Method for Blind Identification of Multichannel FIR Systems in Noise Fields with Unknown Spatial Covariance. *IEEE Signal Processing Letters*, 4(5):135–137, May 1997.

[2] H. Liu, G. Xu, L. Tong, and T. Kailath. Recent Developments in Blind Channel Equalization: From Cyclostationarity to Subspaces. *Signal Processing*, pages 83–99, 1996.

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

[4] P. Van Overschee and B. De Moor. *Subspace Identification for Linear Systems: Theory, Implementation, Applications.* Kluwer Academic Publishers, Dordrecht, 1996.