Blind Subspace Identification of a BPSK Communication Channel

Martin Kristensson* Björn Ottersten* Dirk Slock**
* Dept. of Signals, Sensors & Systems, KTH, S–100 44 Stockholm, SWEDEN
** Institut EURECOM, B.P. 193, F-06904 Sophia Antipolis Cedex, FRANCE

Abstract

This paper considers the problem of blind estimation of multiple FIR channels. When a subspace algorithm is applied to the blind identification problem, incorporating information about the symbol constellation is in general not possible. However, by exploiting special properties of one dimensional symbol constellations (BPSK), it is shown that it is possible to improve or simplify a class of algorithms for blind channel identification. It is also shown that in the case of one dimensional symbol constellations there is a third way, apart from multiple antennas and oversampling, of arriving at a multi channel representation of the communication system.

1 Introduction

Subspace based estimation algorithms have recently been applied to the multi channel identification problem [5]. Oversampled and/or multi receiver signals may be modeled as low rank processes and thus lend themselves to subspace based methods. In [4] a subspace fitting approach is taken to estimate the channel from the covariance matrix. By exploiting the orthogonality property between the noise subspace and the channel matrix, it is shown in [2] that it is possible to identify the channel matrix up to a multiplicative constant. In these algorithms the algebraic structure of the multichannel system is used for identification, but other characteristic properties of communication signals, such as constant modulus or finite alphabet, are not used.

In general, better performance of blind identification algorithms is obtained when properties of the specific symbol constellation are exploited. Combining the algebraic structure of multichannel communication systems with symbol constellation properties is expected to result in algorithms with good performance. In this contribution, we consider the case when the symbol constellation is one dimensional, e.g. binary phase shift keying (BPSK). A generalization of the suggested approach to minimum shift keying (MSK) signaling is also presented. The subspace based methods may in these two specific cases be modified in order to exploit the knowledge of the signal constellation. The proposed algorithms improve the estimation accuracy, lower the computational complexity or make the algorithm insensitive to the noise correlation among the different communication channels. The fact that it is possible to obtain a multi channel representation from one complex communication channel in the case of BPSK or MSK signaling, is also commented on.

2 Data model

Consider a discrete time baseband representation of a digital communication system with linear modulation and L independent linear transmission channels with impulse response length $M+1$ at most. The communication channels may be obtained by sampling an antenna array or by oversampling a single communication channel with respect to the symbol rate. Further on we show a third way to obtain a multichannel representation when the modulation method is either BPSK or MSK.

The received signal in the i^{th} channel can due to the linearity assumption be modeled by

$$x_i(k) = \sum_{l=0}^{M} h_i(l)d(k-l) + n_i(k), \qquad i = 1, \dots, L,$$

where $h_i(l)$ is the complex impulse response of the i^{th} channel and $d(k)$ is the transmitted information symbol at time k. The additive white Gaussian noise $n_i(k)$ is, if not otherwise stated, assumed to be uncorrelated among the channels and equi-powered with variance σ_n^2. Now, study the following vector model of one single communication channel

$$\mathbf{x}_i(k) = \mathbf{H}_i\mathbf{d}(k) + \mathbf{n}_i(k) \qquad i = 1, \dots, L,$$

where the vectors $\mathbf{x}_i(k)$ and $\mathbf{d}(k)$ are defined as

$$\mathbf{x}_i(k) = \begin{bmatrix} x_i(k) & x_i(k-1) & \dots & x_i(k-N+1) \end{bmatrix}^T,$$
$$\mathbf{d}(k) = \begin{bmatrix} d(k) & d(k-1) & \dots & d(k-M-N+1) \end{bmatrix}^T.$$

1058-6393/97 $10.00 © 1997 IEEE

Here N is the width of the temporal window and the $N \times (N+M)$ matrix $\mathbf{H}_i$ is defined as

$$\mathbf{H}_i = \begin{pmatrix} h_i(0) & \dots & h_i(M) & 0 & \dots & 0 \\ 0 & h_i(0) & \dots & h_i(M) & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h_i(0) & \dots & h_i(M) \end{pmatrix}$$

We now have L vector equations describing the communication system. A more convenient vector representation of the system is obtained by collecting the equations

$$\begin{pmatrix} \mathbf{x}_1(k) \\ \vdots \\ \mathbf{x}_L(k) \end{pmatrix} = \begin{pmatrix} \mathbf{H}_1 \\ \vdots \\ \mathbf{H}_L \end{pmatrix} \mathbf{d}(k) + \begin{pmatrix} \mathbf{n}_1(k) \\ \vdots \\ \mathbf{n}_L(k) \end{pmatrix},$$

or with obvious notations

$$\mathbf{x}(k) = \mathbf{H}\mathbf{d}(k) + \mathbf{n}(k). \qquad (1)$$

Note that the noiseless signal in (1) is low rank if the length of the temporal window is chosen large enough to satisfy $NL > N + M$. If data symbols and noise are modeled as zero mean and independent processes, the covariance of the received vector is

$$\mathbf{R}_{xx} = \mathrm{E}\{\mathbf{x}(k)\mathbf{x}^*(k)\} = \mathbf{H}\mathbf{R}_{dd}\mathbf{H}^* + \sigma_n^2\mathbf{I},$$

where $(\cdot)^*$ denotes complex conjugate transpose and $\mathbf{R}_{dd}$ is the covariance of the data vector.

3 General subspace fitting

The low rank structure of the noiseless signal suggests that a subspace based technique may be used for identification of the FIR channels. In [2] a subspace technique using the orthogonality property between the noise and signal subspace is introduced. A statistical analysis of the algorithm can be found in [1].

The specializations for the one dimensional symbol constellation presented in the following two sections are valid for a large class of blind identification algorithms, including the above subspace approach. However, the simulation results presented in the paper are obtained by applying the suggested methods to the subspace algorithm in [2].

4 Using the BPSK assumption

The general subspace method mentioned in the preceding section uses only second order statistics of the received data. Communication signals are in general very structured and the temporal properties, such as constant modulus or finite alphabet, are often used in blind identification/equalization [6]. Below, we show how to exploit a one dimensional symbol constellation (BPSK or PAM in general) to create additional independent channels. This allows the use of a subspace technique based on the second order statistics, still exploiting some of the symbol constellation properties.

By representing the observation vector in terms of real components (rather than an analytic representation) it is possible to exploit the one dimensional symbol constellation. Assume a BPSK communication system with real data symbols and separate (1) into real and imaginary parts

$$\begin{aligned} \tilde{\mathbf{x}}(k) &= \begin{bmatrix} \mathrm{Re}\{\mathbf{x}(k)\} \\ \mathrm{Im}\{\mathbf{x}(k)\} \end{bmatrix} \\ &= \begin{bmatrix} \mathrm{Re}\{\mathbf{H}\} \\ \mathrm{Im}\{\mathbf{H}\} \end{bmatrix} \mathbf{d}(k) + \begin{bmatrix} \mathrm{Re}\{\mathbf{n}(k)\} \\ \mathrm{Im}\{\mathbf{n}(k)\} \end{bmatrix}, \end{aligned}$$

or with obvious notation

$$\tilde{\mathbf{x}}(k) = \tilde{\mathbf{H}}\mathbf{d}(k) + \tilde{\mathbf{n}}(k).$$

The columns of $\tilde{\mathbf{H}}$ span a subspace of the same dimension as $\mathbf{H}$ but the observation space is now real valued of dimension $2NL$. Applying this procedure to one single complex communication channel results in a low rank channel representation. The separation above is thus another way of obtaining a multi channel representation of a communication system. It can obviously be used in combination with oversampling and/or multiple antennas.

In general, as can be seen in the simulations, this separation procedure improves the quality of the estimated channel parameters. If the original channel matrix, $\mathbf{H}$, is full rank, this improvement can be motivated by the following. The original channel matrix has a certain condition number, defined by the quotient of the largest singular value and the smallest nonzero singular value of $\mathbf{H}$. Using some algebraic properties of the new channel matrix, $\tilde{\mathbf{H}}$, the following can be shown,

$$\mathrm{cond}(\tilde{\mathbf{H}}) \leq \mathrm{cond}(\mathbf{H}).$$

The condition number is thus improved (or unchanged) and this should intuitively result in better estimates in the general case. The performance improvement can also be motivated by the fact that $\tilde{\mathbf{H}}$ can be full column rank even if $\mathbf{H}$ is rank deficient.

It has been shown in several papers, e.g. [7], that the channel matrix is full rank if not all the channels share any common zeros. A sufficient condition for the new channel matrix, $\tilde{\mathbf{H}}$, to be full rank, is thus that the real and imaginary parts of the channel do not share any common zeros. A common way to model the taps in a mobile communication channel is to assume that they are Rayleigh fading, i.e. the real and imaginary parts are uncorrelated. This channel model can to some extent justify the assumption that the real and imaginary channel do not share any common zeros.

5 MSK Modulation

Minimum shift keying (MSK), see e.g. [3], is a signaling scheme where the phase of the modulated communication signal is changed with $\pm\pi/2$ at each symbol interval. The phase shift of the signal indicates if the transmitted bit is a zero or a one. Separating the system in exactly the same way as for the BPSK constellation case does in this case not give any additional information for the channel identification. However, if the temporal phase structure of the signal is exploited, similar results are obtained. To see this, again study the matrix formulation of the communication system

$$\mathbf{x}(k) = \mathbf{H}\mathbf{d}(k) + \mathbf{n}(k).$$

The symbols $d(k)$ belong to the alphabet $\{\pm 1\}$ at even symbol periods and to the alphabet $\{\pm i\}$ at odd symbol periods. In fact, the symbols $d(k)$ can be interpreted as a modulated version of the symbols $b(k)$, i.e. $d(k) = i^k b(k)$, where the symbols $b(k)$ belong to the alphabet $\{\pm 1\}$. Now introduce a diagonal matrix in which the m^{th} diagonal element is equal to i^{m-1}, i.e.,

$$\mathbf{J} = \text{diag}(1, i, -1, -i, 1, \dots),$$

and modulated versions thereof, $\mathbf{J}_k = i^{-k}\mathbf{J}$. Using these matrices we get

$$\begin{aligned}\mathbf{x}(k) &= \mathbf{H}\mathbf{J}_k^{-1}(\mathbf{J}_k\mathbf{d}(k)) + \mathbf{n}(k)\\ &= (\mathbf{H}\mathbf{J}_k^{-1})\mathbf{b}(k) + \mathbf{n}(k)\end{aligned}$$

which shows that we are close to the BPSK case. However, the matrices $(\mathbf{H}\mathbf{J}_k^{-1})$ are time-varying and as a result, the vector process $\mathbf{x}(k)$ is not stationary but cyclostationary, which hampers the time-averaging process for obtaining covariance estimates. However, by down-modulating the received vectors $\mathbf{x}(k)$, we get

$$\begin{aligned}\mathbf{z}(k) = i^{-k}\mathbf{x}(k) &= (\mathbf{H}\mathbf{J}^{-1})\mathbf{b}(k) + i^{-k}\mathbf{n}(k)\\ &= (\mathbf{H}\mathbf{J}^{-1})(\mathbf{J}_k\mathbf{d}(k)) + i^{-k}\mathbf{n}(k)\end{aligned}$$

where $\mathbf{z}(k)$ is now a stationary vector process. The MSK communication system is thus equivalent to a BPSK system where the channel matrix and the input signals are replaced by $\mathbf{H}\mathbf{J}^{-1}$ and $i^{-k}\mathbf{J}\mathbf{d}(k)$ respectively. Taking the real and imaginary parts of the model gives

$$\begin{aligned}\tilde{\mathbf{z}}(k) &= \begin{bmatrix}\text{Re}\{\mathbf{z}(k)\}\\ \text{Im}\{\mathbf{z}(k)\}\end{bmatrix}\\ &= \begin{bmatrix}\text{Re}\{\mathbf{H}\mathbf{J}^{-1}\}\\ \text{Im}\{\mathbf{H}\mathbf{J}^{-1}\}\end{bmatrix}(i^{-k}\mathbf{J}\mathbf{d}(k)) + \begin{bmatrix}\text{Re}\{i^{-k}\mathbf{n}(k)\}\\ \text{Im}\{i^{-k}\mathbf{n}(k)\}\end{bmatrix}.\end{aligned}$$

Note that the channel matrix is now low rank but has a slightly different structure compared to (1). However, comparing this new channel matrix with the original one suggests that also the new matrix is identifiable from its column space, this even if the original channel is only one dimensional.

6 Alternative subspace fitting

Another way to exploit the one dimensional symbol constellation property is to note that the complex Gaussian noise is circularly symmetric but the data symbols are not. This can be formulated mathematically by

$$\begin{aligned}&E\{\mathbf{n}(k)\mathbf{n}^T(k)\} = 0,\\ &E\{\mathbf{d}(k)\mathbf{d}^*(k)\} = \text{E}\{\mathbf{d}(k)\mathbf{d}^T(k)\} = \mathbf{R}_{dd},\end{aligned}$$

where $(\cdot)^T$ denotes transpose. Now, study the expected value of $\mathbf{x}(k)\mathbf{x}^T(k)$,

$$\mathbf{R}_{xx^c} = E\{\mathbf{x}(k)\mathbf{x}^T(k)\} = \mathbf{H}\mathbf{R}_{dd}\mathbf{H}^T, \tag{2}$$

where $(\cdot)^c$ denotes complex conjugate. In this way the noise contribution to the covariance expression disappears but the channel part is still present. A subspace method based on $\mathbf{R}_{xx^c}$ is thus, as long as the noise terms have a circularly symmetric distribution, insensitive to the noise correlation structure. If the multichannel representation is obtained by oversampling with respect to the symbol rate, a correlation between the noise samples is probable, but using the suggested approach in this section, unbiased identification of the channel coefficients is still possible. This is not the case for a subspace fitting approach based on $\mathbf{R}_{xx}$. In the simulation we will see how the general subspace fitting method applied to $\mathbf{R}_{xx^c}$ instead of $\mathbf{R}_{xx}$ is more robust with respect to the additive noise properties.

Within this alternative approach, one particular computationally simple channel estimate can be obtained as follows. Observe from (2) that the column spaces of $\mathbf{H}$ and of $\mathbf{R}_{xx^c}$ are the same. This leads us to introduce the following subspace fitting criterion

$$\min_{\mathbf{h},\mathbf{Q}} \left\|\mathbf{H} - \hat{\mathbf{R}}_{xx^c}\mathbf{B}\mathbf{Q}\right\|_F^2, \tag{3}$$

where $||.||_F$ denotes Frobenius norm and $\mathbf{h}$ is a vector containing the channel parameters. The matrix $\mathbf{B}$ has the same dimensions as $\mathbf{H}$ and is fixed. Its choice influences the quality of the channel estimate. The criterion is separable in $\mathbf{h}$ and $\mathbf{Q}$. Minimizing w.r.t. $\mathbf{Q}$ first yields

$$\mathbf{Q} = (\mathbf{F}^*\mathbf{F})^{-1}\mathbf{F}^*\mathbf{H}\,, \quad \mathbf{F} = \hat{\mathbf{R}}_{xx^c}\mathbf{B}\,.$$

Substitution in (3) yields

$$\min_{\mathbf{h}} \left\|\mathbf{P}_{\mathbf{F}}^{\perp}\mathbf{H}\right\|_F^2,$$

where $\mathbf{P}_{\mathbf{F}}^{\perp} = \mathbf{I} - \mathbf{P}_{\mathbf{F}}$ and $\mathbf{P}_{\mathbf{F}} = \mathbf{F}(\mathbf{F}^*\mathbf{F})^{-1}\mathbf{F}^*$. This criterion has to be minimized subject to a non-triviality constraint on $\mathbf{h}$. If we choose $||\mathbf{h}|| = 1$, then we get

$$\hat{\mathbf{h}} = \arg\max_{||\mathbf{h}||=1} \text{Tr}\{\mathbf{H}^*\mathbf{P}_{\mathbf{F}}\mathbf{H}\} = \arg\max_{||\mathbf{h}||=1} \mathbf{h}^*\mathcal{F}\mathbf{h},$$

where $\mathcal{F}$ can easily be constructed from $\mathbf{P_F}$. The solution for $\mathbf{h}$ is hence the eigenvector of $\mathcal{F}$ corresponding to its maximum eigenvalue. As far as the choice of $\mathbf{B}$ is concerned, it appears that $\mathbf{B} = \mathbf{H}^c$ would be optimal, considering the structure of $\mathbf{R}_{xx^c}$. Simulation results confirm that this choice of $\mathbf{B}$ is optimal or very close to optimal. Hence a two-step procedure suggests itself: solve the above problem twice, once with $\mathbf{B}$ an arbitrary selection matrix, yielding a consistent estimate $\hat{\mathbf{h}}$. Then solve the problem a second time with $\mathbf{B} = \hat{\mathbf{H}}^c$ in which the channel estimate from the first step, $\hat{\mathbf{h}}$ is used. In the simulation runs the selection matrix, $\mathbf{B}$, in the first step is chosen as the first columns of the identity matrix. This low complexity method differs from the other approaches in that no eigendecomposition of the complete covariance matrix is necessary. The only eigendecomposition necessary is the computation of one eigenvector of a matrix which is of the same dimension as the total channel length $(M+1)L$. In the following simulation section this last method is referred to as the low complexity algorithm.

7 Simulation results

The suggested methods have been applied to the subspace algorithm based on the orthogonality property and tested in simulation runs. In all simulations the signal to noise ratio is defined according to

$$SNR = 10\log_{10}\frac{\mathrm{E}\{\|\mathbf{Hd}(k)\|^2\}}{\mathrm{E}\{\|\mathbf{n}(k)\|^2\}},$$

and the estimation error covariance according to

$$COV = 10\log_{10}\frac{\mathrm{E}\{\|\hat{\mathbf{H}} - \mathbf{H}\|^2\}}{\|\mathbf{H}\|^2}.$$

In the above definitions $\mathbf{H}$ and $\hat{\mathbf{H}}$ are the channel matrices obtained with the temporal window length, N, equal to 1. The expected value in the covariance definition of the channel estimates is replaced by the sample average in the simulations. The presented results are throughout this section the averages over 2000 different noise realizations.

In the first simulation setup the BPSK and the MSK approaches were applied to the same single communication channel with the channel coefficients equal to

$$[1.0+1.0i, 0.45+0.00i, -0.70+1.20i,$$
$$0.42+0.13i, -0.32+0.58i].$$

The length of the temporal window was chosen to $N = 6$ and 500 symbols were used for the estimation of the covariance matrix. The results are shown in Figure 1. The simulation indicates that in both the case of BPSK signals and MSK signals, it is possible to identify the single complex communication channel with an algebraic approach for reasonable signal to noise ratios. However, the results vary with the transmission channel. Sometimes the case of BPSK signaling works better and sometimes MSK signaling. With only one complex communication channel available, the identification method is very sensitive to the placement of channel zeros in the complex plane. This sensitivity agrees well with the identifiability conditions.

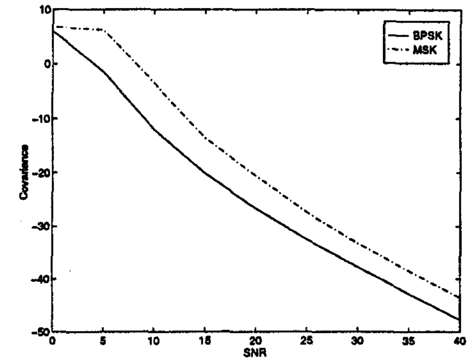

Figure 1. Identification of a single complex channel.

Next we study the case of a multichannel (4 channels) communication link. The signaling on the channel is in this case of BPSK type. The simulation compares the general subspace fitting approach with the suggested algorithms in this paper. The order of the four channels is $M = 4$, the length of the temporal window $N = 3$ and 500 symbols were used for the estimation of the covariance matrix. The

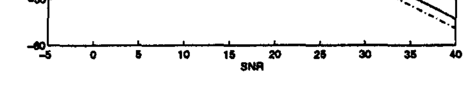

Figure 2. Identification of four channels.

general subspace fitting approach is marked with a solid line, the low complexity algorithm as described in section 6 with a dashed line, the separation approach in section 4 with a dashed-dotted line and finally the subspace fitting approach where $\mathbf{R}_{xx^c}$ is used instead of $\mathbf{R}_{xx}$ is marked with a dotted line. The simulation shows that the method where the received data is separated into a real and imaginary part outperforms the other methods for all signal to noise ratios. All the other methods behave approximately the same. Note that the low complexity algorithm behaves equally well as the more complicated general subspace fitting method. For the low complexity algorithm to perform well, the second step in the algorithm is essential. To illustrate this, the accuracy of the estimates obtained in the first and second step of the low complexity algorithm is compared in Figure 3.

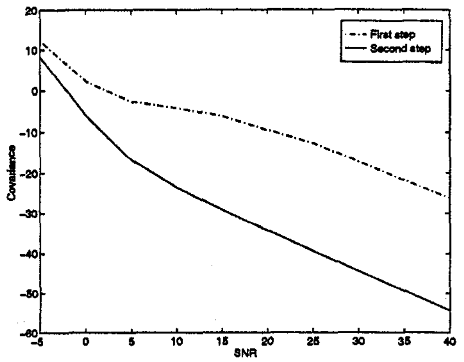

Figure 3. Low complexity algorithm, first and second steps.

Finally we demonstrate the fact that if the noise is correlated among the channels, the methods using the circularity property of the noise will be the only ones that have unbiased estimates. Except for the noise the same simulation setup is used as in the previous case. The noise in the four channels has the following structure of the covariance sequence,

$$\mathrm{E}\{\mathbf{n}(k)\mathbf{n}^*(l)\} = \sigma_n^2 \begin{bmatrix} 1 & 0.5 & 0 & 0 \\ 0.5 & 1 & 0.5 & 0 \\ 0 & 0.5 & 1 & 0.5 \\ 0 & 0 & 0.5 & 1 \end{bmatrix} \delta(k-l).$$

The simulation agrees well with intuition. For small signal to noise ratios, i.e. where the noise structure affects the estimates significantly, the circular approach is the best alternative to use. As the signal to noise ratio improves, the circular approach coincides with the general approach and the separation approach behaves the best.

8 Conclusions

In this paper the symbol constellation properties are used to improve the quality of the channel estimates for a large class of blind identification algorithms. In addition we show that in the case of a one dimensional symbol constellation, it is possible to obtain a multichannel representation of the communication system using only one complex communication channel. This enables the use of efficient algorithms using only the algebraic structure of the second order statistics also for one dimensional systems. A generalization of the method to MSK signals is also presented. Finally, in the case of circular noise, a low complexity alternative subspace fitting method has been presented.

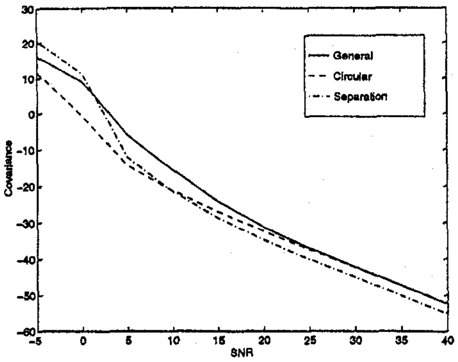

Figure 4. Correlated noise.

References

[1] M. Kristensson and B. Ottersten. Statistical analysis of a subspace method for blind channel identification. In *Proc. ICASSP*, volume 5, pages 2437–2440, Atlanta, Georgia, May 1996. IEEE.

[2] E. Moulines, P. Duhamel, J.-F. 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.

[3] J. G. Proakis. *Digital Communications*. McGraw-Hill, third edition, 1995.

[4] D. T. Slock and C. B. Papadias. Blind fractionally-spaced equalization based on cyclostationarity. In *Vehicular Technology Conference*, volume 2, pages 1286–1290. IEEE, June 1994.

[5] L. Tong, G. Xu, and T. Kailath. A new approach to blind identification and equalization of multipath channels. In *Proc. of the 25th Asilomar Conference on Signals, Systems and Computers*, pages 856–860, Pacific Grove, CA, November 1991.

[6] J. Treichler and B. Agee. A new approach to multipath correction of constant modulus signals. *IEEE Trans. Acoust. Speech, Signal Processing*, 31:459–472, April 1983.

[7] J. K. Tugnait. On blind identifiability of multipath channels using fractional sampling and second-order cyclostationary statistics. *IEEE Transactions on Information Theory*, 41(1):308–311, January 1995.