LINEAR PREDICTION AND SUBSPACE FITTING BLIND CHANNEL IDENTIFICATION BASED ON CYCLIC STATISTICS.

Luc Deneire and Dirk T.M. Slock

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

presented at: "13th international Conference on Digital Signal Processing", Santorini, Greece, July 1997.



LINEAR PREDICTION AND SUBSPACE FITTING BLIND CHANNEL IDENTIFICATION BASED ON CYCLIC STATISTICS.

Luc Deneire¹ and Dirk T.M. Slock

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

Abstract : Blind channel identification and equalization based on second-order statistics by subspace fitting and linear prediction have received a lot of attention lately. On the other hand, the use of cyclic statistics in fractionally sampled channels has also raised considerable interest. We propose to use these statistics in subspace fitting and linear prediction for (possibly multiuser and multiple antennas) channel identification. We base our identification schemes on the cyclic statistics, using the stationary multivariate representation introduced by [2] and [4] [5]. This leads to the use of all cyclic statistics. The methods proposed appear to have good performance.

1. PROBLEM POSITION

We consider a communication system with p emitters and a receiver constituted of an array of M antennas. The signals received are oversampled by a factor m w.r.t. the symbol rate. The channel is FIR of duration NT/m where T is the symbol duration. The received signal can be written as :

$$\mathbf{x}(n) = \sum_{\substack{k=-\infty \\ k=-\infty}}^{\infty} \mathbf{h}(k)\mathbf{u}(n-k) + \mathbf{v}(n)$$
$$= \sum_{\substack{k=-\infty \\ k=-\infty}}^{\infty} \mathbf{h}(n-km)\mathbf{a}_k + \mathbf{v}(n)$$

where

$$\mathbf{u}(n) = \sum_{k=-\infty}^{\infty} \mathbf{a}_k \delta(n-km)$$

The received signal $\mathbf{x}(n)$ and noise $\mathbf{v}(n)$ are a $M \times 1$ vectors. $\mathbf{x}(n)$ is cyclostationary with period m whereas $\mathbf{v}(n)$ is assumed not to be cyclostationary with period m. $\mathbf{h}(k)$ has dimension $M \times p$, $\mathbf{a}(k)$ and $\mathbf{u}(k)$ have dimensions $p \times 1$.

2. CYCLIC STATISTICS

Following the assumptions hereabove, the correlations :

$$\mathbf{R}_{xx}(n,\tau) = \mathbf{E}\left\{\mathbf{x}(n)\mathbf{x}^{H}(n-\tau)\right\}$$

are cyclic in n with period m (^H denotes complex conjugate tranpose). One can easily express them as:

$$\mathbf{R}_{xx}(n,\tau) = \sum_{\alpha=-\infty}^{\infty} \sum_{\beta=-\infty}^{\infty} \mathbf{h}(n-\alpha m) \mathbf{R}_{aa}(\beta) \mathbf{h}^{H}(n-\alpha m+\beta m-\tau) + \mathbf{R}_{vv}(\tau)$$

We then express the k^{th} cyclic correlation as :

$$\mathbf{R}_{xx}^{\{k\}}(\tau) \stackrel{\Delta}{=} \frac{1}{m} \sum_{l=0}^{m-1} \mathbf{R}_{xx}(l,\tau) e^{-j\frac{2\pi lk}{m}}$$
$$= \mathbf{E}^{k} \left\{ \mathbf{x}(l) \mathbf{x}^{H}(l-\tau) \right\}$$

whose value is :

$$\mathbf{R}_{xx}^{\{k\}}(\tau) = \frac{1}{m} \sum_{\alpha=-\infty}^{\infty} \sum_{\beta=-\infty}^{\infty} \mathbf{h}(\alpha) \mathbf{R}_{aa}(\beta) \\ \mathbf{h}^{H}(\alpha + \beta m - \tau) e^{-j\frac{2\pi\alpha k}{m}} \\ + \mathbf{R}_{vv}(\tau) \delta(k)$$

We can introduce a cyclic correlation matrix as :

$$\mathbf{R}_{xx}^{\{k\}} \triangleq \begin{bmatrix} \mathbf{R}_{xx}^{\{k\}}(0) & \mathbf{R}_{xx}^{\{k\}}(1) & \cdots & \mathbf{R}_{xx}^{\{k\}}(K-1) \\ \mathbf{R}_{xx}^{\{k\}}(-1) & \mathbf{R}_{xx}^{\{k\}}(0) & \cdots & \mathbf{R}_{xx}^{\{k\}}(K-2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{xx}^{\{k\}}(1-K) & \mathbf{R}_{xx}^{\{k\}}(2-K) & \cdots & \mathbf{R}_{xx}^{\{k\}}(0) \end{bmatrix} = \mathcal{T}_{K}(\mathbf{H}_{N}\mathbf{D}_{DFT}^{\{k\}})\mathbf{R}_{uu}^{\{k\}}\mathcal{T}_{K}^{H}(\mathbf{H}_{N}) + \delta(k)\mathbf{R}_{vv}$$

where $\mathbf{R}_{uu}^{\{k\}} = \mathbf{R}_{aa} \otimes I_m$ and \otimes is a block Kronecker product, the first matrix is a block matrix and the second matrix is an elementwise matrix.

 $\mathcal{T}_{K}(\mathbf{H}_{N})$ is the convolution matrix of $\mathbf{H}_{N} = [\mathbf{h}(0)^{T}\mathbf{h}(1)^{T}\cdots\mathbf{h}(N-1)^{T}]^{T}$ and

$$\mathbf{D}_{DFT}^{\{k,p\}} = \mathsf{blockdiag}[\mathbf{I}_p | e^{-j\frac{2\pi k}{m}} \mathbf{I}_p | \cdots | e^{-j\frac{2\pi (N-1)k}{m}} \mathbf{I}_p]$$

3. GLADYSHEV'S THEOREM AND MIAMEE PROCESS

Gladyshev's theorem [2] states that :

Theorem 1 Function $\mathbf{R}_{xx}(n, \tau)$ is the correlation function of some PCS (Periodically Correlated Sequence) iff the matrix-valued function :

$$\underline{\mathbf{R}}(\tau) = \left[\mathbf{R}_{xx}^{\{kk'\}}(\tau)\right]_{k,k'=0}^{m-1}$$

where
$$\mathbf{R}_{xx}^{\{kk'\}}(\tau) = \mathbf{R}_{xx}^{\{k-k'\}}(\tau)e^{2\pi jk\tau/m}$$

is the matricial correlation function of some m-variate stationary sequence.

¹The work of Luc Deneire is supported by the EC by a Marie-Curie Fellowship (TMR program) under contract No ERBFMBICT950155

Reminding that $\mathbf{R}_{xx}^{\{k\}}(\tau) = \mathbf{R}_{xx}^{\{m-k\}H}(-\tau)$, the following matrix

$$\mathbf{\underline{R}} \stackrel{\Delta}{=} \begin{bmatrix} \mathbf{\underline{R}}(0) & \mathbf{\underline{R}}(1) & \cdots & \mathbf{\underline{R}}(K-1) \\ \mathbf{\underline{R}}(-1) & \mathbf{\underline{R}}(0) & \cdots & \mathbf{\underline{R}}(K-2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{\underline{R}}(1-K) & \mathbf{\underline{R}}(2-K) & \cdots & \mathbf{\underline{R}}(0) \end{bmatrix}$$

is an hermitian $K \times K$ block Toeplitz matrix of $Mm \times Mm$ blocks.

Then, Miamee [4] gives us the explicit expression of the multivariate stationary process associated :

$$\underline{\mathbf{Z}}_{n} = \left[\mathbf{Z}_{n}^{k}\right]_{k=0}^{m-1} \text{ where } \mathbf{Z}_{n}^{k} = \bigoplus_{\substack{j=0\\j=0}}^{m-1} \mathbf{x}(n+j)e^{2\pi jk(n+j)/n}$$

where \oplus is the direct sum, i.e., noting $w = e^{2\pi j/m}$

$$\mathbf{Z}_n^k = w^{kn}[\mathbf{x}(n), \mathbf{x}(n+1)w^k, \cdots, \mathbf{x}(n+m-1)w^{k(m-1)}]$$

is defined in a Hilbert space, where the correlation is the following euclidian product :

$$<\mathbf{Z}_{n}^{k},\mathbf{Z}_{n+l}^{k'}>=\sum_{j=0}^{m-1}\mathbb{E}\left\{\mathbf{Z}_{n}^{k}(j)\mathbf{Z}_{n+l}^{k'}{}^{H}(j)\right\}$$

and $\underline{\mathbf{Z}}_n = [\mathbf{Z}_n^{0^T} \mathbf{Z}_n^{1^T} \cdots \mathbf{Z}_n^{m-1^T}]^T$ with the classical correlation for multivariate stationary processes.

On the other hand, Miamee gives the link between the linear prediction on $\underline{\mathbf{Z}}_n$ and the cyclic AR model of $\mathbf{x}(n)$.

4. EXPRESSION OF \underline{Z}_n w.r.t. $\mathbf{u}(n)$ and $\mathbf{h}(n)$

From
$$\mathbf{Z}_{n}^{k} = \bigoplus_{j=0}^{m-1} \mathbf{x}(n+j)e^{2\pi j k(n+j)/m}$$
 and
 $\mathbf{x}(n+j) = \sum_{k=0}^{L-1} \mathbf{h}(k)\mathbf{u}(n+j-k) + \mathbf{v}(n+j)$

$$= \mathbf{H}_{N} \begin{bmatrix} \mathbf{u}(n+j) \\ \mathbf{u}(n+j+1) \\ \vdots \\ \mathbf{u}(n+j-N+1) \end{bmatrix}$$

$$+ \mathbf{v}(n+j)$$

Defining $\mathbf{U}_{n+j} = [\mathbf{u}(n+j)^T \cdots \mathbf{u}(n+j-N+1)^T]^T$ and $\mathbf{H}_N^{\{k\}} = [w^{-kj}\mathbf{h}(j)]_{j=0}^{N-1}$ we express the Miamee process as :

$$\begin{aligned} \mathbf{Z}_{n}^{k} &= \bigoplus_{\substack{j=0\\j=0}}^{m-1} (\mathbf{H}_{N}^{\{-k\}} w^{kn} \mathbf{U}_{n+j} + \mathbf{v}(n+j) e^{2\pi j k \frac{n+j}{m}}) \\ &= \mathbf{H}_{N}^{\{-k\}} w^{kn} [\mathbf{U}_{n} \mathbf{U}_{n+1} \cdots \mathbf{U}_{n+m-1}] \\ &+ \bigoplus_{j=0}^{m-1} \mathbf{v}(n+j) e^{2\pi j k \frac{n+j}{m}} \end{aligned}$$

$$\Rightarrow \underline{\mathbf{Z}}_{n} = \mathbf{H}_{tot} \mathcal{U}(n) + \mathcal{V}(n)$$

where we noted $\mathbf{H}_{tot} = [\mathbf{H}_{N}^{\{0\}T} \mathbf{H}_{N}^{\{-1\}T} \cdots \mathbf{H}_{N}^{\{1-m\}T}]^{T}$

$$\mathcal{U}(n) = \mathbf{D}_{DFT}^{\{n,pN\}} [\mathbf{U}_n \mathbf{U}_{n+1} \cdots \mathbf{U}_{n+m-1}]$$

and $\mathcal{V}(n) =$
$$\begin{bmatrix} \mathbf{v}(n) & \cdots & \mathbf{v}(n+m-1) \\ \mathbf{v}(n)w^n & \cdots & \mathbf{v}(n+m-1)w^{n+m-1} \\ \vdots & \ddots & \vdots \\ \mathbf{v}(n)w^{n(m-1)} & \cdots & \mathbf{v}(n+m-1)w^{(m-1)(n+m-1)} \\ \Rightarrow \mathbf{\underline{Z}} = \mathcal{T}_{L+N-1}(\mathbf{H}_{tot})\mathcal{U}_L + \mathcal{V}_L$$
(1)

where $\mathcal{U}_L = [\mathcal{U}(n)]_{n=L-1}^0$ clearly is a stationary process whose correlation matrix can easily be deduced from \mathbf{R}_{aa} .

Based on relation (1), we apply the classical subspace fitting and linear prediction channel identification schemes, as detailed below.

5. SIGNAL SUBSPACE FITTING

We recall briefly the signal subspace fitting (noise subspace based) blind channel identification algorithm hereunder.

One can write the (compact form of the) SVD of the cyclocorrelation matrix $\underline{\mathbf{R}} = \mathbf{U}\mathbf{D}\mathbf{V}^H$ with the relations:

range{**U**} = range{**V**} = range{
$$\mathcal{T}_K(\mathbf{H}_{tot})$$
}

We have assumed that $\mathcal{T}_K(\mathbf{H}_{tot})$ is full rank, which leads to the usual identifiability condition. We can then solve the classical subspace fitting problem :

$$\min_{\mathbf{H}_{tot},T} ||\mathcal{T}_K(\mathbf{H}_{tot}) - \mathbf{U}T||_F^2$$

If we introduce \mathbf{U}^{\perp} such that $[\mathbf{U}\mathbf{U}^{\perp}]$ is a unitary matrix, this leads to

$$\min_{\mathbf{H}_{tot}} \mathbf{H}_{tot}^{t} \left[\sum_{i=D^{\perp}}^{KMm} \mathcal{T}_{N}(\mathbf{U}_{i}^{\perp Ht}) \mathcal{T}_{N}^{H}(\mathbf{U}_{i}^{\perp Ht}) \right] \mathbf{H}_{tot}^{Ht}$$

where \mathbf{U}_i^{\perp} is a $KMm^2 \times 1$, $D^{\perp} = N + K$ and superscript ^t denotes the transposition of the blocks of a block matrix. Under constraint $||\mathbf{H}_{tot}|| = 1$, $\hat{\mathbf{H}}_{tot}^t$ is then the eigenvector corresponding to the minimum eigenvalue of the matrix between brackets. One can lower the computational burden by using $D^{\perp} > N + K$ (see a.o. [6]).

The case p > 1 can be (partially) solved in a manner similar to [7] and [3].

6. LINEAR PREDICTION

We consider the denoised case. The correlation matrix is then computed as follows.

$$\mathbf{R}_{xx,sb}^{\{0\}} = \mathbf{R}_{xx}^{\{0\}} - \mathbf{R}_{\mathcal{V}\mathcal{V}}(\tau)$$
 yields :

$$\begin{aligned} \left[\mathbf{R}_{\mathcal{V}\mathcal{V}}(\tau) \right]_{i,j} \\ &= \sum_{l=0}^{m-1} \mathbf{E} \left\{ \mathbf{v}(n+l) \mathbf{v}^H(n+l+\tau) \right\} w^{i(n+l)} w^{-j(n+l+\tau)} \\ &= \mathbf{R}_{vv}(\tau) w^{n(i-j)-j\tau} \sum_{l=0}^{m-1} w^{(i-j)l} \\ &= \mathbf{R}_{vv}(\tau) w^{n(i-j)-j\tau} m \delta_{ij} \\ &= m \delta_{ij} \mathbf{R}_{vv}(\tau) w^{-j\tau} \end{aligned}$$

Hence $\mathbf{R}_{\mathcal{V}\mathcal{V}}(\tau) = \mathbf{R}_{vv}(\tau)$ blockdiag $[\mathbf{I}_M | w^{\tau} \mathbf{I}_M | w^{2\tau} \mathbf{I}_M | \cdots | w^{(m-1)\tau} \mathbf{I}_M]$, which, in $\underline{\mathbf{R}}$, corresponds to the noise contribution of the zero cyclic frequency cyclic correlation.

From equation (1) and noting $\mathcal{Z}_K(n-1) = [\underline{\mathbf{Z}}_j]_{j=n-1}^{j=n-K}$, the predicted quantities are :

$$\hat{\underline{Z}}(n)|_{\mathcal{Z}_{K}(n-1)} = \mathbf{p}_{1}\underline{\mathbf{Z}}_{n-1} + \dots + \mathbf{p}_{K}\underline{\mathbf{Z}}_{n-K}$$

$$\tilde{\underline{Z}}(n) = \underline{\mathbf{Z}}(n) - \hat{\underline{\mathbf{Z}}}(n)|_{\mathcal{Z}_{K}(n-1)}$$

Following [9], we rewrite the correlation matrix as

$$\underline{\mathbf{R}} = \left[\begin{array}{cc} \mathbf{R}_o & \mathbf{r}_K \\ \mathbf{r}_K^H & \mathbf{R}_{K-1} \end{array} \right]$$

this yields the prediction filter :

$$\mathbf{P}_K \stackrel{\Delta}{=} [\mathbf{p}_1 \cdots \mathbf{p}_K] = -\mathbf{r}_K \mathbf{R}_{K-1}^{-1}$$

and the prediction error variance :

$$\sigma_{\underline{\tilde{\mathbf{Z}}},K} = \mathbf{R}_o - \mathbf{P}_K \mathbf{r}_K^H$$

where the inverse might be replaced by the Moore-Penrose pseudoinverse, and still yield a consistent channel estimate. Another way of being robust to order overestimation would be to use the Levinson-Wiggins-Robinson (LWR) algorithm to find the prediction quantities and estimate the order with this algorithm.

Lots of ways are possible to go from the prediction quantites to the channel estimate ([8] and [1]).

For our purpose, we used the simple suboptimal solution hereunder.

From the prediction error equation, it is easy to derive :

$$\mathbf{H}_{tot}^{Tt} \mathcal{T}_{K}([\mathbf{I}_{Mm} \mathbf{P}_{K}^{t}]) = [\mathbf{H}_{tot}^{Tt} \mathbf{0}]$$

Hence $\mathbf{I}_{MmK} - \mathbf{H}_{tot}^{Tt} \mathcal{T}_K ([\mathbf{I}_{Mm} \mathbf{P}_K^t]) = [\mathbf{H}_{tot}^{Tt}(0)\mathbf{0}]$; and \mathbf{H}_{tot} is found by minimizing $\mathbf{I}_{Mm(K-1)} - \mathbf{H}_{tot}^{Tt} \mathcal{T}_K (\mathbf{P}_K^t)$ and is thus the left singular vector corresponding to the minimum singular value of this matrix. This solution corresponds to a "plain least-squares" solution, and is robust w.r.t. order overestimation. This also means that it is not the best solution available, but this discussion is beyond the scope of this paper.

7. COMPUTATIONAL ASPECTS

It is obvious that the correlation matrix $\underline{\mathbf{R}}$ built from the cyclic correlations is bigger (in fact each scalar in \mathbf{R} is replaced by a $m \times m$ block in $\underline{\mathbf{R}}$) than the corresponding matrix built from the classical Time Series representation of oversampled stationary signals. This fact must be balanced with the stronger structure that is cast in our correlation matrix. In fact, one can (not so easily) prove that the estimates $\hat{\mathbf{H}}_N^{\{-k\}}$ are strictly related (i.e. $\hat{\mathbf{H}}_N^{\{-k\}} = [w^{-kj}\hat{\mathbf{h}}(j)]_{j=0}^{N-1}$ for all k), which indicates us that this structure should lead to reduced complexity algorithms w.r.t. the original ones. When developing the expressions in detail, this is particularly obvious in linear prediction, where the prediction filter has some strong structure (which is also visible in [5]).

8. SIMULATIONS

In our simulations, we restrict ourselves to the p = 1 case, using a randomly generated real channel of length 6T, an oversampling factor of m = 3 and M = 3 antennas. We draw the NRMSE of the channel, defined as

$$\mathbf{NRMSE} = \sqrt{\frac{1}{100} \sum_{l=1}^{100} ||\hat{\mathbf{h}}^{(l)} - \mathbf{h}||_{F}^{2} / ||\mathbf{h}||_{F}^{2}}$$

where $\hat{\mathbf{h}}^{(l)}$ is the estimated channel in the l^{th} trial. In the figures below, the NRMSE in dB has been calculated as $20 * \log 10(NRMSE)$.

The correlation matrix is calculated from a burst of 100 QAM-4 symbols (note that if we used real sources, we would have used the conjugate cyclocorrelation, wich is another means of getting rid of the noise, provided it is circular). For these simulations, we used 100 Monte-Carlo runs.

8.1. Subspace fitting

The estimations of 25 realisations, for an SNR of 20 dB, are reproduced hereunder.



For comparison, we used the same algorithm for the classical Time Series representation of the oversampled signal. The results hereunder show a better performance for the classic approach, which is due to the fact that we used the same complexity for both algorithms (same matrix size), which results in a lower noise subspace size for the cyclic approach. In theory, when one uses the same subspace size, as there is a one for one correspondance between the elements of the classic correlation matrix and the elements of the cyclic correlation matrix, the performances should be equal. The third curve illustrates this fact.



8.2. Linear prediction

For the linear prediction, we expect to have a slightly better performance in the cyclic approach then in the classic approach. Indeed, in the classic approach, if we use for example M = 1 antenna and an oversampling factor of m = 3, we predict $[x(n)x(n-1)x(n-2)]^T$ based on $[x(n-3)x(n-4)\cdots]^T$, whereas in the cyclic approach we predict the scalar x(n) based on $[x(n-1)x(n-2)x(n-3)\cdots]^T$. The corresponding prediction filter thus captures little more prediction features in the cyclic case.

On the other hand, the noise contribution being only present in the zero cyclic frequency cyclic correlation, we expect a better behavior of the method if we don't take the noise into account in the correlation matrix (i.e. we don't estimate the noise variance before doing the linear prediction). Those expectations are confirmed by the following simulations, note that the mention LP on cyclic statistics refers to the use of \mathbf{R} where the noise contribution has been removed, whereas the mention LP on cyclic statistics, no "denoising" refers to the use of the plain correlation matrix.



9. CONCLUSIONS

Using the stationary multivariate representation introduced by [2] and [4] [5], we have explicitly expressed this process. It can be seen as the ouput of a system with transfer channel $\mathbf{H}_{tot} = [\mathbf{H}_N^{\{0\}T} \mathbf{H}_N^{\{-1\}T} \cdots \mathbf{H}_N^{\{1-m\}T}]^T$ and input easily related to the actual system input. Once these quantities expressed, application of the classical subspace fitting and linear prediction algorithms is straightforward.

For the subspace fitting, one has essentially the same performance as in the Time Series Representation [9]. The only advantage one could expect is some refinement in the channel order estimation prior to the subspace fitting. The main drawback is the increase of the computational burden. For the linear prediction, we get a better performance due to the fact that we take the very near past into account. Although the complexity is, for the moment, far more heavy. Use of the LWR algorithm and of the structure of the correlation matrix should lead to similar calculation loads.

Further work on these topics should stress on reducing the complexity of these algorithms.

References

- A. Ghorokov and Ph. Loubaton. "On blind identification of certain multivariate MA models". *submitted to IEEE Transactions on Signal Processing*, 1996.
- [2] E.G. Gladyshev. "Periodically random sequences". *Soviet Mathematics*, 2, 1961.
- [3] Ph. Loubaton. "Quelques aspects méthodologiques concernant la séparation des mélanges convolutifs".
 "Ecole des Techniques avancées Signal Image Parole, Grenoble", septembre 1996.
- [4] A.G. Miamee. "Periodically Correlated Processes and their Stationary Dilations". SIAM Journal of Applied Mathematics, 50(4):1194–1199, August 1990.
- [5] A.G. Miamee. "Explicit Formula for the Best Linear Predictor of Periodically Correlated Sequences". *SIAM Journal of Mathematic Analysis*, 24(3):703– 711, May 1993.
- [6] 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.
- [7] Dirk T.M. Slock. "Blind Joint Equalization of Multiple Synchronous Mobile Users Using Oversampling and/or Multiple Antennas". In *Twenty-Eight Asilomar Conference on Signal, Systems & Computers*, October 1994.
- [8] Dirk T.M. Slock and Constantinos B. Papadias. "Further Results on Blind Identification and Equalization of Multiple FIR Channels". In "Proc. ICASSP 95, Detroit, Michigan, May 1995", May 1995.
- [9] D.T.M. Slock and C.B. Papadias. "Blind Fractionally-Spaced Equalization Based on Cyclostationarity". In *Proc. Vehicular Technology Conf.*, pages 1286–1290, Stockholm, Sweden, June 1994.