BLIND FRACTIONALLY-SPACED EQUALIZATION, PERFECT-RECONSTRUCTION FILTER BANKS AND MULTICHANNEL LINEAR PREDICTION

Dirk T.M. Slock

Eurecom Institute, 2229 route des Crêtes
B.P. 193, 06904 Sophia Antipolis Cedex, FRANCE

ABSTRACT

Equalization for digital communications constitutes a very particular blind deconvolution problem in that the received signal is cyclostationary. Oversampling (OS) (w.r.t. the symbol rate) of the cyclostationary received signal leads to a stationary vector-valued signal (polyphase representation (PR)). OS also leads to a fractionally-spaced channel model and equalizer. In the PR, channel and equalizer can be considered as an analysis and synthesis filter bank. Zero-forcing (ZF) equalization corresponds to a perfect-reconstruction filter bank. We show that in the OS case FIR ZF equalizers exist for a FIR channel. In the PR, the multichannel linear prediction of the noiseless received signal becomes singular eventually, reminiscent of the single-channel prediction of a sum of sinusoids. As a result, the channel can be identified from the received signal second-order statistics by linear prediction in the noise-free case, and by using the Pisarenko method when there is additive noise. In the given data case, Music (subspace) or ML techniques can be applied.

1. PREVIOUS WORK

Consider linear digital modulation over a linear channel with additive Gaussian noise so that the received signal can be written as

$$y(t) \;=\; \sum_k a_k h(t-kT) \;+\; v(t) \tag{1}$$

where the a_k are the transmitted symbols, T is the symbol period, $h(t)$ is the combined impulse response of channel and transmitter and receiver filters, but is often called the channel response for simplicity. Assuming the $\{a_k\}$ and $\{v(t)\}$ to be (wide-sense) stationary, the process $\{y(t)\}$ is (wide-sense) cyclostationary with period T. If the channel would be known, then one could pass the received signal through a matched filter and sample the output at the symbol rate. These samples would provide sufficient statistics for the detection of the transmitted symbols. If $\{y(t)\}$ is sampled with period T, the sampled process is (wide-sense) stationary and its second-order statistics contain no information about the phase of the channel. Tong, Xu and Kailath [1] have proposed to oversample the received signal with a period $\Delta = T/m$, $m > 1$. In what follows, we assume $h(t)$ to have a finite duration. Tong *et al.* have shown that the channel can be identified from the second-order statistics of the oversampled received signal. They introduce an observation vector $\mathbf{y}(k)$ of received samples over a certain time window and consider a matrix linear model of the form

$$\mathbf{y}(k) \;=\; \mathbf{H}\,\mathbf{a}(k) + \mathbf{v}(k)\,. \tag{2}$$

The drawback of their approach is that they need the sampled channel matrix $\mathbf{H}$ to have full column rank. This leads to an unnecessary overparameterization of the channel as will become clear below (the matrix $\mathbf{H}$ could be parameterized in terms of the samples of the channel response, but this parameterization is not exploited by Tong *et al.*). Tong *et al.* found that the condition for identifiability of the (oversampled) channel from the second-order statistics of the received signal is that the z-transform of the oversampled channel should not have m equispaced zeros on a circle centered in the origin. One should also remark that the identification of the channel from the received signal second-order statistics can only be done up to a multiplicative constant (with magnitude one in certain cases), a not unusual phenomenon in blind equalization. This constant can be identified by other means. If the channel contains a delay, then this delay can also not be identified blindly. We shall consider here an oversampling factor $m = 2$, but the results can be generalized [2].

2. FRACTIONALLY-SPACED CHANNELS AND EQUALIZERS, AND FILTER BANKS

We assume the channel to be FIR with duration NT. We consider the polyphase description of the received signal. With $m = 2$, let $y_1(k)$ and $y_2(k)$ denote the even and odd samples of $y(t)$ ($y_1(k) = y(t_0+kT)$, $y_2(k) = y(t_0+(k-\frac{1}{2})T)$), and similarly for the noise samples and channel response. Then the oversampled received signal can be represented in vector form at the symbol rate as

$$\begin{aligned}
&\mathbf{y}(k) \;=\; \sum_{i=0}^{N-1} \mathbf{h}(i)a_{k-i} + \mathbf{v}(k) \;=\; \mathbf{H}_N A_N(k) + \mathbf{v}(k)\,,\\
&\mathbf{y}(k) = \begin{bmatrix} y_1(k) \\ y_2(k) \end{bmatrix},\ \mathbf{v}(k) = \begin{bmatrix} v_1(k) \\ v_2(k) \end{bmatrix},\ \mathbf{h}(k) = \begin{bmatrix} h_1(k) \\ h_2(k) \end{bmatrix},\\
&\mathbf{H}_N = [\mathbf{h}(0) \cdots \mathbf{h}(N-1)]\,,\ \ A_N(k) = \left[a_k^H \cdots a_{k-N+1}^H\right]^H\,,
\end{aligned} \tag{3}$$

where superscript ^{H} denotes Hermitian transpose. We formalize the finite duration NT (approximately) assumption of the channel as follows

$$(\text{AFIR}):\ \ \mathbf{h}(0) \neq 0,\ \mathbf{h}(N-1) \neq 0,\ \mathbf{h}(i) = 0 \text{ for } i \geq N.$$

The z-transform of the channel response at the sampling rate is $H(z) = H_1(z^2) + z^{-1}H_2(z^2)$. Similarly, consider a fractionally-spaced ($\frac{T}{2}$) equalizer of which the z-transform can also be decomposed into its polyphase components: $F(z) = F_1(z^2) + z^{-1}F_2(z^2)$, see Fig. 1. As will become clear below, a unique ZF and properly scaled equalizer can be found (under certain conditions on the channel) when F_1 and F_2 are FIR filters of length $N-1$. However, in light of the prediction and noise subspace parameterization considerations to be discussed further, we take F_1, F_2 to be FIR of length N: $F_i(z) = \sum_{k=0}^{N-1} f_i(k)z^{-k}$, $i = 1, 2$.

0-7803-1775-0/94 $3.00 © 1994 IEEE

3. FIR ZERO-FORCING (ZF) EQUALIZATION

The condition for the equalizer to be ZF is $F_1(z)H_1(z) + F_2(z)H_2(z) = 1$. If we introduce $\mathbf{f}(k) = [f_1(k)\ f_2(k)]$, $\mathbf{F}_N = [\mathbf{f}(0)\cdots\mathbf{f}(N-1)]$, then the ZF condition can be written as

$$\mathbf{F}_N\ \mathcal{T}_N(\mathbf{H}_N) = [1\ 0\cdots 0] \tag{4}$$

where $\mathcal{T}_M(\mathbf{x})$ is a (block) Toeplitz matrix with M (block) rows and $\left[\mathbf{x}\ 0_{p\times(M-1)}\right]$ as first (block) row (p is the number of rows in $\mathbf{x}$). (4) is a system of $2N-1$ equations in $2N$ unknowns. The equalizer has one degree of freedom more than necessary to be zero-forcing. Let us arbitrarily constrain $f_2(0) = 0$. Furthermore, consider equalization with removal of ISI, but up to a constant only and take $f_1(0) = 1$. Then the remaining equalizer coefficients can be found from

$$[\mathbf{f}(1)\cdots\mathbf{f}(N-1)]\ \mathcal{T}_{N-1}(\mathbf{H}_N) = -[h_1(1)\cdots h_1(N-1)\ 0\cdots 0] \tag{5}$$

where $\mathcal{T}_{N-1}(\mathbf{H}_N)$ is now a square matrix of size $2N-2$. $\mathcal{T}_{N-1}(\mathbf{H}_N)$ is a Sylvester matrix (up to a permutation) which is known to be nonsingular if $H_1(z)$ and $H_2(z)$ have no zeros in common. This condition coincides with the identifiability condition of Tong *et al.* on $H(z)$. Let us denote the resulting equalizer coefficients as $\mathbf{F}_N^{p1}$. A set of equalizer coefficients that satisfies (4) is $\mathbf{F}_N = \mathbf{F}_N^{p1}/h_1(0)$ (assuming $h_1(0) \neq 0$). The unique ZF equalizer of length $N-1$ is

$$\mathbf{F}_{N-1} = [1\ 0\cdots 0]\ \mathcal{T}_{N-1}^{-1}(\mathbf{H}_N)\ . \tag{6}$$

Note that there exists a set of blocking equalizer coefficients $\mathbf{F}_N^b$ for which no transmitted symbol has an influence on the equalizer output:

$$\mathbf{F}_N^b \mathcal{T}_N(\mathbf{H}_N) = 0 \tag{7}$$

(the nullspace of $\mathcal{T}_N^H(\mathbf{H}_N)$ has dimension one).

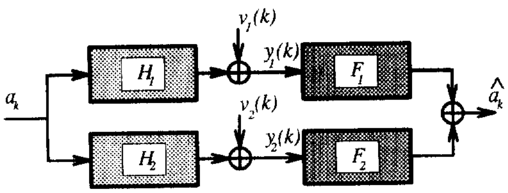

Figure 1. Polyphase representation of the T/2 fractionally-spaced channel and equalizer.

4. CHANNEL IDENTIFICATION FROM SECOND-ORDER STATISTICS BY MULTICHANNEL LINEAR PREDICTION

In this section, we consider the noiseless case: $v(t) \equiv 0$. Similarly to $\mathbf{F}^{p1}$, we can introduce $\mathbf{F}^{p2} = [0\ 1\ *\cdots *]$ so that $\mathbf{F}^{p1}$, $\mathbf{F}^{p2}$ satisfy

$$\begin{bmatrix} \mathbf{F}_N^{p1} \\ \mathbf{F}_N^{p2} \end{bmatrix} \mathcal{T}_N(\mathbf{H}_N) = \begin{bmatrix} h_1(0) \\ h_2(0) \end{bmatrix} [1\ 0\cdots 0]\ . \tag{8}$$

Consider now the problem of predicting $\mathbf{y}(k)$ from $\mathbf{Y}_{N-1}(k-1) = \left[\mathbf{y}^H(k-1)\cdots\mathbf{y}^H(k-N+1)\right]^H$. The prediction error can be written as

$$\tilde{\mathbf{y}}(k)|_{\mathbf{Y}_{N-1}(k-1)} = \mathbf{y}(k) - \hat{\mathbf{y}}(k)|_{\mathbf{Y}_{N-1}(k-1)} = [I_2\ \ -\mathbf{P}_{N-1}]\ \mathbf{Y}_N(k)\ . \tag{9}$$

Minimizing the prediction error variance leads to the following optimization problem

$$\min_{\mathbf{P}_{N-1}}\ [I_2\ \ -\mathbf{P}_{N-1}]\,\mathrm{R}_N^{\mathbf{Y}}\,[I_2\ \ -\mathbf{P}_{N-1}]^H = \sigma^2_{\tilde{\mathbf{y}},N-1} \tag{10}$$

where $\mathrm{R}_M^a = \mathrm{E}\ A_M(k)A_M^H(k)$ and

$$\mathrm{R}_N^{\mathbf{Y}} = \mathrm{E}\ \mathbf{Y}_N(k)\mathbf{Y}_N^H(k) = \mathcal{T}_N(\mathbf{H}_N)\,\mathrm{R}_{2N-1}^a\,\mathcal{T}_N^H(\mathbf{H}_N)\ . \tag{11}$$

Exploiting (11) in (10), and assuming that $H_1(z)$ and $H_2(z)$ have no common zeros, one can show that $\mathbf{P}_{N-1}$ satisfies

$$[I_2\ \ -\mathbf{P}_{N-1}]\,\mathcal{T}_N(\mathbf{H}_N)\begin{bmatrix} \mathbf{Q}_{2N-2} \\ I_{2N-2} \end{bmatrix} = 0\ , \tag{12}$$

where $\mathbf{Q}_{2N-2}$ are the least-squares predictor coefficients for

$$\hat{a}(k)|_{A_{2N-2}(k-1)} = \mathbf{Q}_{2N-2}\ A_{2N-2}(k-1)\ , \tag{13}$$

$$\left[1\ \ -\mathbf{Q}_{2N-2}\right]\mathrm{R}_{2N-1}^a = \left[\sigma^2_{\tilde{a},2N-2}\ 0\cdots 0\right]\ . \tag{14}$$

Since the orthogonal complement of the rightmost matrix in (12) has dimension one, (12) leads to

$$[I_2\ \ -\mathbf{P}_{N-1}]\,\mathcal{T}_N(\mathbf{H}_N) = \mathbf{h}(0)\left[1\ \ -\mathbf{Q}_{2N-2}\right] \tag{15}$$

which, by postmultiplication with $A_{2N-1}(k)$, can be translated to

$$\tilde{\mathbf{y}}(k)|_{\mathbf{Y}_{N-1}(k-1)} = \mathbf{h}(0)\ \tilde{a}(k)|_{A_{2N-2}(k-1)}\ . \tag{16}$$

Using (16), one can show

$$\sigma^2_{\tilde{\mathbf{y}},N-1} = \sigma^2_{\tilde{a},2N-2}\,\mathbf{h}(0)\mathbf{h}^H(0)\ . \tag{17}$$

By comparing (8) and (15), we see that when the transmitted data are uncorrelated ($\mathrm{R}_{2N-1}^a = \sigma_a^2 I_{2N-1}$, $\mathbf{Q}_{2N-2} = 0$), we have

$$\begin{bmatrix} \mathbf{F}_N^{p1} \\ \mathbf{F}_N^{p2} \end{bmatrix} = [I_2\ \ -\mathbf{P}_{N-1}]\ . \tag{18}$$

If the transmitted symbols are correlated however, the prediction filter $\mathbf{P}_{N-1}$ is affected. Not completely however, since

$$[-h_2(0)\ \ h_1(0)]\,[I_2\ \ -\mathbf{P}_{N-1}] = [-h_2(0)\ \ h_1(0)]\begin{bmatrix} \mathbf{F}_N^{p1} \\ \mathbf{F}_N^{p2} \end{bmatrix} \tag{19}$$

which is proportional to $\mathbf{F}_N^b$ and only depends on $\mathbf{H}_N$. Note that $[-h_2(0)\ \ h_1(0)]$ can be determined up to a scalar multiple from the variance expression in (17) so that $\mathbf{F}_N^b$ can be determined from $\mathbf{P}_{N-1}$ and the variance expression, and hence from the prediction problem.

Let us introduce a block-componentwise transposition operator ^{t}, viz.

$$\begin{aligned} \mathbf{H}_N^t &= [\mathbf{h}(0)\cdots\mathbf{h}(N-1)]^t = \left[\mathbf{h}^T(0)\cdots\mathbf{h}^T(N-1)\right] \\ \mathbf{F}_N^t &= [\mathbf{f}(0)\cdots\mathbf{f}(N-1)]^t = \left[\mathbf{f}^T(0)\cdots\mathbf{f}^T(N-1)\right] \end{aligned} \tag{20}$$

where ^{T} is the usual transposition operator. Now the channel can be identified from the blocking equalizer. Indeed, (7) leads to $\sum_{i=1}^2 F_i^b(z)\,H_i(z) = 0$ which is also

$$\mathbf{F}_N^b\ \mathcal{T}_N(\mathbf{H}_N) = 0 \iff \mathbf{H}_N^t\ \mathcal{T}_N\left(\mathbf{F}_N^{bt}\right) = 0\ . \tag{21}$$

This last equation allows one to determine the channel coefficients $\mathbf{H}_N$ up to a constant, if again $H_1(z)$ and $H_2(z)$ have no zeros in common (which is the same condition as $F_1^n(z)$ and $F_2^n(z)$ having no zeros in common since $F_1^b(z)/F_2^b(z) = -H_2(z)/H_1(z)$). A set of unique coefficients can be obtained by introducing one extra constraint. Traditionally, two types of constraints have been considered for this purpose. A quadratic constraint: $\left\|\mathbf{H}_N^t\right\| = 1$. In this case, $\mathbf{H}_N^{tH}$ is the $2N^{\text{th}}$ left singular vector of $\mathcal{T}_N\left(\mathbf{F}_N^{bt}\right)$. Or a linear constraint: $\mathbf{H}_N^t\, g = 1$. In this case $\mathbf{H}_N^{tH}$ is proportional to the last column of the Q matrix in the unnormalized QR factorization of the matrix $\left[\mathcal{T}_N\left(\mathbf{F}_N^{bt}\right)\ \ g\right]$. The first approach is numerically more reliable. If the symbol variance σ_a^2 is known, then from the prediction error variance expression in (17), we can identify $|h_1(0)|$ (or $|h_2(0)|$ if $h_1(0) = 0$). So we have identified the channel from the received signal second-order statistics, up to a factor $\frac{h_1(0)}{|h_1(0)|}$.

To recapitulate, in the absence of additive noise, we have a singular prediction problem. From the multichannel prediction error variance and the prediction coefficients, one can identify the null space of the covariance matrix, the blocking equalizer $\mathbf{F}_N^b$. From $\mathbf{F}_N^b$, one can identify the channel up to multiplicative constant as indicated above. From (15), one can identify $\mathbf{Q}_{2N-2}$ and via (14), this leads to the identification of the (Toeplitz) symbol covariance matrix R_{2N-1}^a up to the multiplicative scalar σ_a^2 (which may be known). If the transmitted symbols are uncorrelated, then the prediction problem immediately provides ZF equalizers, see (18),(8). If the transmitted symbols are correlated, then a FIR ZF equalizer can still be found directly from the FIR channel. The ZF equalizer with shortest length is given in (6).

5. SIGNAL AND NOISE SUBSPACES

Suppose now that we have additive white noise $v(t)$ with zero mean and unknown variance σ_v^2 (in the complex case, real and imaginary parts are assumed to be uncorrelated, colored noise could equally well be handled [2]). Then since

$$\mathrm{R}_N^{\mathbf{y}} = \mathcal{T}_N(\mathbf{H}_N)\,\mathrm{R}_{2N-1}^a\mathcal{T}_N^H(\mathbf{H}_N) + \sigma_v^2 I_{2N}\,, \qquad (22)$$

σ_v^2 can be identified as the smallest eigenvalue of $\mathrm{R}_N^{\mathbf{y}}$, and the corresponding eigenvector is $\mathbf{F}_N^b$. This is the Pisarenko method [3, page 500]. By replacing $\mathrm{R}_N^{\mathbf{y}}$ by $\mathrm{R}_N^{\mathbf{y}} - \sigma_v^2 I_{2N}$, all results of the prediction approach in the noiseless case still hold.

Consider now a covariance matrix of size $M \geq N$. Given $\mathrm{R}_N^{\mathbf{y}}$, we have been able to identify all the desired quantities in the case $M = N$. So given covariance information, there cannot be anything to be gained from considering $M > N$. However, this is not necessarily the case when the covariance sequence is estimated from data. So consider the block Toeplitz matrix $\mathcal{T}_M(\mathbf{H}_N)$ of dimension $2M \times (M+N-1)$. The following lemma is easy to show.

Lemma 1 *With assumption (AFIR) and assuming that $H_1(z)$ and $H_2(z)$ are coprime,*

$$\textit{rank}\ (\mathcal{T}_M(\mathbf{H}_N)) = M+N-1\,, \quad M \geq N.$$

Hence, under the assumptions of the lemma, $\mathcal{T}_M(\mathbf{H}_N)$ has full column rank. The orthogonal complement of the space spanned by the columns of $\mathcal{T}_M(\mathbf{H}_N)$ therefore has dimension $M-N+1$. With the blocking equalizer $\mathbf{F}_N^b$ satisfying $\mathbf{F}_N^b\ \mathcal{T}_N(\mathbf{H}_N) = 0$, it is easy to see that

$$\mathcal{T}_{M-N+1}\left(\mathbf{F}_N^b\right)\ \mathcal{T}_M(\mathbf{H}_N) = 0\,, \quad M \geq N \qquad (23)$$

where $\mathcal{T}_{M-N+1}\left(\mathbf{F}_N^b\right)$ is a $(M-N+1) \times 2M$ block Toeplitz matrix in which the blocks are 1×2. Under the conditions of the lemma above, $\mathcal{T}_{M-N+1}\left(\mathbf{F}_N^b\right)$ has full (row) rank. Hence, the columns of $\mathcal{T}_{M-N+1}^H\left(\mathbf{F}_N^b\right)$ span the orthogonal complement of the column space of $\mathcal{T}_M(\mathbf{H}_N)$. Given the structure of

$$\mathrm{R}_M^{\mathbf{y}} = \mathcal{T}_M(\mathbf{H}_N)\,\mathrm{R}_{M+N-1}^a\mathcal{T}_M^H(\mathbf{H}_N) + \sigma_v^2 I_{2M}\,, \qquad (24)$$

the column spaces of $\mathcal{T}_M(\mathbf{H}_N)$ and $\mathcal{T}_{M-N+1}^H\left(\mathbf{F}_N^b\right)$ are called the signal and noise subspaces respectively.

Consider the eigendecomposition of $\mathrm{R}_M^{\mathbf{y}}$ of which the real nonnegative eigenvalues are ordered in descending order:

$$\mathrm{R}_M^{\mathbf{y}} = \sum_{i=1}^{M+N-1} \lambda_i V_i V_i^H + \sum_{i=M+N}^{2M} \lambda_i V_i V_i^H = V_S\Lambda_S V_S^H + V_{\mathcal{N}}\Lambda_{\mathcal{N}} V_{\mathcal{N}}^H \qquad (25)$$

where $\Lambda_{\mathcal{N}} = \sigma_v^2 I_{M-N+1}$ (see (24)). Assuming $\mathcal{T}_M(\mathbf{H}_N)$ and R_{M+N-1}^a to have full rank, the sets of eigenvectors V_S and $V_{\mathcal{N}}$ are orthogonal: $V_S^H V_{\mathcal{N}} = 0$, and $\lambda_i > \sigma_v^2$, $i = 1,\ldots,M+N-1$. We then have the following equivalent descriptions of the signal and noise subspaces

$$\begin{aligned} \textit{Range}\ \{V_S\} &= \textit{Range}\ \{\mathcal{T}_M(\mathbf{H}_N)\} \\ \textit{Range}\ \{V_{\mathcal{N}}\} &= \textit{Range}\ \left\{\mathcal{T}_{M-N+1}^H\left(\mathbf{F}_N^b\right)\right\}\ . \end{aligned} \qquad (26)$$

In particular,

$$V_{\mathcal{N}}^H\mathcal{T}_M(\mathbf{H}_N) = 0\,, \quad \mathcal{T}_{M-N+1}\left(\mathbf{F}_N^b\right)\ V_S = 0\,. \qquad (27)$$

6. CHANNEL ESTIMATION FROM AN ESTIMATED COVARIANCE SEQUENCE BY SUBSPACE FITTING

When the covariance matrix is estimated from data, it will no longer satisfy exactly the properties we have elaborated upon. A first (detection) problem then is to determine the dimension of the signal subspace. A number of techniques for doing this have been elaborated in the literature (typically based on an investigation of the eigenvalues) and we shall assume that the correct dimension $M+N-1$ (and hence the correct channel order N) has been detected. We shall again order the eigenvalues and eigenvectors as in (25). The signal subspace will now be defined as the space spanned by the eigenvectors corresponding to the $M+N-1$ largest eigenvalues, while the noise subspace corresponds to the $M-N+1$ remaining eigenvectors (as in (25), except that $\Lambda_{\mathcal{N}}$ is no longer a multiple of the identity matrix).

Consider now the following subspace fitting problem

$$\min_{\mathbf{H}_N,T} \|\mathcal{T}_M(\mathbf{H}_N) - V_S T\|_F \qquad (28)$$

where the Frobenius norm of a matrix Z can be defined in terms of the trace operator: $\|Z\|_F^2 = \operatorname{tr}\left\{Z^H Z\right\}$. The problem considered in (28) is quadratic in both $\mathbf{H}_N$ and T. If V_S contains the signal subspace eigenvectors of the actual covariance matrix $\mathrm{R}_M^{\mathbf{y}}$, then the minimal value of the cost function in (28) is zero. Indeed, if the column spaces of two matrices with full column rank are identical (as in (26)), then one of the matrices can be transformed into the other one by postmultiplication with a unique nonsingular square matrix. If $\mathrm{R}_M^{\mathbf{y}}$ is estimated from a finite amount of data however, then its eigenvectors (and eigenvalues) are perturbed w.r.t. their theoretical values. Therefore, in general there will be no value for $\mathbf{H}_N$ for which the column space of $\mathcal{T}_M(\mathbf{H}_N)$ coincides with the signal subspace *Range* $\{V_S\}$.

But it is clearly meaningful to try to estimate $\mathbf{H}_N$ by taking that $\mathcal{T}_M(\mathbf{H}_N)$ into which V_S can be transformed with minimal cost. This leads to the subspace fitting problem in (28). The optimization problem in (28) is separable. With $\mathbf{H}_N$ fixed, the optimal matrix T can be found to be (assuming $V_S^H V_S = I$)

$$T = V_S^H \mathcal{T}_M(\mathbf{H}_N) \,. \tag{29}$$

Using (29) and the commutativity of the convolution operator as in (21), one can show that (28) is equivalent to

$$\begin{aligned}&\min_{\mathbf{H}_N^t} \mathbf{H}_N^t \left(\sum_{i=M+N}^{2M} \mathcal{T}_M\left(V_i^{H\,t}\right) \mathcal{T}_M^H\left(V_i^{H\,t}\right) \right) \mathbf{H}_N^{t\,H} \\ &= \min_{\mathbf{H}_N^t} \left[M \left\|\mathbf{H}_N^t\right\|_2^2 - \mathbf{H}_N^t \left(\sum_{i=1}^{M+N-1} \mathcal{T}_M\left(V_i^{H\,t}\right) \mathcal{T}_M^H\left(V_i^{H\,t}\right) \right) \mathbf{H}_N^{t\,H} \right]\end{aligned} \tag{30}$$

where V_i^H (like $\mathbf{F}_N$) is considered a block vector with M blocks of size 1×2. These optimization problems have to be augmented with a nontriviality constraint on $\mathbf{H}_N^t$. In case we choose the quadratic constraint $\left\|\mathbf{H}_N^t\right\|_2 = 1$, then the last term in (30) leads equivalently to

$$\max_{\left\|\mathbf{H}_N^t\right\|_2 = 1} \mathbf{H}_N^t \left(\sum_{i=1}^{M+N-1} \mathcal{T}_M\left(V_i^{H\,t}\right) \mathcal{T}_M^H\left(V_i^{H\,t}\right) \right) \mathbf{H}_N^{t\,H} \tag{31}$$

the solution of which is the eigenvector corresponding to the maximum eigenvalue of the matrix appearing between the brackets. In the case of $m = 2$, the noise subspace always has a lower dimension than the signal subspace. Hence it is computationally more interesting to estimate $\mathcal{T}_M(\mathbf{H}_N)$ by optimizing its orthogonality to the noise subspace, rather than by optimizing its fit to the signal subspace.

Alternatively, we may consider the following subspace fitting problem

$$\min_{\mathbf{F}_N^b, T} \left\| \mathcal{T}_{M-N+1}^H\left(\mathbf{F}_N^b\right) - V_N T \right\|_F \tag{32}$$

which leads again to either a minimization problem optimizing the orthogonality to the signal subspace, or a maximization problem optimizing the fit to the noise subspace. In this case, the latter will be computationally more interesting. The channel $\mathbf{H}_N$ can then be identified from $\mathbf{F}_N^b$ as we discussed before.

When $M = N$, the subspace fitting problem in (28) leads to the Pisarenko method discussed before. When $M > N$, the Pisarenko method generalizes to the Music method [3, page502] (corresponding to (30)). When the exact covariance matrix is given, any value of $M \geq N$ will lead to the same value for $\mathbf{H}_N$, namely the true channel (up to a multiplicative scalar). When the covariance matrix is estimated from data, the estimated covariance lags can be considered as a noisy version of the true ones and hence a better estimate should be obtained as more data are incorporated, as M increases. However, as M increases, the quality of the covariance matrix estimate from a fixed finite amount of data goes down. So there should be some optimal value for M, compromising for these two opposite effects.

7. CHANNEL AND TRANSMITTED SYMBOLS ESTIMATION FROM DATA USING DETERMINISTIC ML

In the case of given data (samples of $y(.)$), the subspace fitting approach of the previous section involves the data through the sample covariance matrix. Though this leads to computationally tractable optimization problems, this may not lead to very efficient estimates from an estimation theoretic point of view. Therefore we consider here a deterministic or conditional maximum likelihood (DML) method. The likelihood is conditional on the transmitted symbols and the channel parameters, which are hence treated as deterministic unknowns. The stochastic part only comes from the additive noise, which we shall assume Gaussian and white with zero mean and unknown variance σ_v^2 ($\overline{\mathrm{R}}_{2M}^v = I_{2M}$, though the generalization to any known $\overline{\mathrm{R}}_{2M}^v$ is straightforward). We assume the data $\mathbf{Y}_M(k)$ to be available. The maximization of the likelihood function boils down to the following least-squares problem

$$\min_{\mathbf{H}_N, A_{M+N-1}(k)} \left\|\mathbf{Y}_M(k) - \mathcal{T}_M(\mathbf{H}_N) A_{M+N-1}(k)\right\|_2^2 \,. \tag{33}$$

The optimization problem in (33) is again separable. Eliminating $A_{M+N-1}(k)$ in terms of $\mathbf{H}_N$, we get

$$\min_{\mathbf{H}_N} \left\| P_{\mathcal{T}_M(\mathbf{H}_N)}^{\perp} \mathbf{Y}_M(k) \right\|_2^2 \,. \tag{34}$$

Now we can use the equivalent parameterization through $\mathbf{F}_N^b$ of the orthogonal complement of *Range* $\{\mathcal{T}_M(\mathbf{H}_N)\}$ to obtain

$$\min_{\mathbf{H}_N} \left\| P_{\mathcal{T}_M(\mathbf{H}_N)}^{\perp} \mathbf{Y}_M(k) \right\|_2^2 = \min_{\mathbf{F}_N^b} \left\| P_{\mathcal{T}_{M-N+1}^H\left(\mathbf{F}_N^b\right)} \mathbf{Y}_M(k) \right\|_2^2 \tag{35}$$

Because of the commutativity of convolution, we can again rewrite

$$\mathcal{T}_{M-N+1}\left(\mathbf{F}_N^b\right) \mathbf{Y}_M(k) = \mathcal{H}_{M-N+1}\left(\mathbf{Y}_M^t(k)\right) \mathbf{F}_N^{b\,T} \tag{36}$$

where $\mathcal{H}_L(\mathbf{x})$ is a block Hankel matrix with L block rows, obtained by taking the block entries from the block vector $\mathbf{x}$ and filling up a Hankel matrix starting from the top left corner. (36) allows us to rewrite the criterion (35) as

$$\begin{aligned}&\min_{\mathbf{F}_N^b} \mathbf{F}_N^{b\,*} \mathcal{H}_{M-N+1}^H\left(\mathbf{Y}_M^t(k)\right) \\ &\left(\mathcal{T}_{M-N+1}\left(\mathbf{F}_N^b\right) \mathcal{T}_{M-N+1}^H\left(\mathbf{F}_N^b\right)\right)^{-1} \mathcal{H}_{M-N+1}\left(\mathbf{Y}_M^t(k)\right) \mathbf{F}_N^{b\,T}\end{aligned} \tag{37}$$

where $\mathbf{F}_N^{b\,*} = \mathbf{F}_N^{b\,TH}$. Again, this criterion has to be augmented with a nontriviality constraint. The optimization problem in (37) is nonlinear. It can easily be solved iteratively in such a way that in each iteration, a quadratic problem appears [4]. An initial estimate may be obtained from the subspace fitting approach discussed above. See [2] for a discrete stochastic ML approach.

REFERENCES

[1] 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 & Computers*, pages 856–860, Pacific Grove, CA, Nov. 1991.

[2] Dirk T.M. Slock. "Blind Fractionally-Spaced Equalization and Channel Identification from Second-Order Statistics or Data". Technical Report RR N° 93-002, Institut Eurécom, Sophia Antipolis, France, 17 Aug. 1993.

[3] L.L. Scharf. *Statistical Signal Processing*. Addison-Wesley, Reading, MA, 1991.

[4] Y. Bresler and A. Macovski. "Exact Maximum Likelihood Parameter Estimation of Superimposed Exponential Signals in Noise". *IEEE Trans. Acoust., Speech and Sig. Proc.*, ASSP-34:1081–1089, Oct. 1986.