FURTHER RESULTS ON BLIND IDENTIFICATION AND EQUALIZATION OF MULTIPLE FIR CHANNELS

Dirk T. M. Slock and Constantinos B. Papadias

Institut EURECOM
2229 route des Crêtes, B.P. 193
06904, Sophia Antipolis Cedex, FRANCE
email: {slock,papadias}@eurecom.fr

ABSTRACT

In previous work, we have shown that in the case of multiple antennas and/or oversampling, FIR ZF equalizers exist for FIR channels and can be obtained from the noise-free linear prediction (LP) problem. The LP problem also lead to a minimal parameterization of the noise subspace, which was used to solve the deterministic maximum likelihood (DML) channel estimation problem. Here we present further contributions along two lines. One is a number of blind equalization techniques of the adaptive filtering type. We also present some robustifying modifications of the DML problem.

1. INTRODUCTION

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 (overall) channel impulse response. The cyclostationarity of $\{y(t)\}$ and the fact that after sampling, multiple received signals from multiple antennas and/or oversampling lead to a vector of received samples at the symbol rate have been discussed in [1],[2],[3],[4]. The case of multiple synchronous transmitting sources has been treated in [5], but here we'll stick to one source. We assume the channel to be FIR with duration of approximately NT. The vector received signal at the symbol rate can be written (for m symbol-rate discrete-time channels) as

$$\mathbf{y}(k) \;=\; \sum_{i=0}^{N-1} \mathbf{h}(i)a_{k-i} + \mathbf{v}(k) \;=\; \mathbf{H}_N A_N(k) + \mathbf{v}(k)\,,$$
$$\mathbf{y}(k) = \begin{bmatrix} y_1(k) \\ \vdots \\ y_m(k) \end{bmatrix}, \mathbf{v}(k) = \begin{bmatrix} v_1(k) \\ \vdots \\ v_m(k) \end{bmatrix}, \mathbf{h}(k) = \begin{bmatrix} h_1(k) \\ \vdots \\ h_m(k) \end{bmatrix}$$
$$\mathbf{H}_N = [\mathbf{h}(0)\cdots\mathbf{h}(N-1)]\,, A_N(k) = \left[a_k^H \cdots a_{k-N+1}^H\right]^H \tag{2}$$

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

$$\mathbf{h}(0) \neq 0,\; \mathbf{h}(N-1) \neq 0 \text{ and } \mathbf{h}(i) = 0 \text{ for } i < 0 \text{ or } i \geq N. \tag{3}$$

We introduce a multichannel equalizer structure as in Fig. 1, i.e., consider a set of m FIR filters of length L operating on the m received signals and take the sum of the filter outputs as the equalizer output. Let the equalizer coefficients be $\mathbf{f}(k) = [f_1(k)\cdots f_m(k)]$, $\mathbf{F}_L = [\mathbf{f}(0)\cdots\mathbf{f}(L-1)]$, and consider the channel and equalizer transfer functions $H(z) = \sum_{k=0}^{N-1}\mathbf{h}(k)z^{-k}$ and $F(z) = \sum_{k=0}^{L-1}\mathbf{f}(k)z^{-k}$. The condition for the equalizer to be zero-forcing (ZF) up to some delay n is $F(z)H(z) = z^{-n}$ where $n = 0, 1, \ldots, N+L-2$. The ZF condition can be written in the time-domain as

$$\mathbf{F}^{ZF}_{L,n}\; \mathcal{T}_L(\mathbf{H}_N) \;=\; [0\cdots0\;\;1\;\;0\cdots0] \tag{4}$$

where the 1 is in the $n{+}1^{st}$ position and $\mathcal{T}_M(\mathbf{x})$ 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 $L{+}N{-}1$ equations in Lm unknowns. To be able to equalize, we need to choose the equalizer length L such that the system of equations (4) is exactly or underdetermined. Hence

$$L \;\geq\; \underline{L} \;=\; \left\lceil \frac{N-1}{m-1} \right\rceil\;. \tag{5}$$

The matrix $\mathcal{T}_L(\mathbf{H}_N)$ is a generalized Sylvester matrix. It can be shown that for $L \geq \underline{L}$ it has full column rank if the FIR assumption (3) is satisfied, and if $H(z) \neq 0,\ \forall z$ or in words if the different channel responses have no zeros in common. Assuming $\mathcal{T}_L(\mathbf{H}_N)$ to have full column rank, the nullspace of $\mathcal{T}_L^H(\mathbf{H}_N)$ has dimension $L(m{-}1){-}N{+}1$. If we take the entries of any vector in this nullspace as equalizer coefficients, then the equalizer output is zero, regardless of the transmitted symbols.

Consider now the measured data with additive independent white noise $\mathbf{v}(k)$ and assume $\mathrm{E}\mathbf{v}(k)\mathbf{v}^H(k) = \sigma_v^2 I_m$. A vector of L measured data can be expressed as

$$\mathbf{Y}_L(k) \;=\; \mathcal{T}_L(\mathbf{H}_N)\; A_{L+N-1}(k) + \mathbf{V}_L(k) \tag{6}$$

where $\mathbf{Y}_L(k) = \left[\mathbf{y}^H(k)\cdots\mathbf{y}^H(k{-}L{+}1)\right]^H$ and $\mathbf{V}_L(k)$ is defined similarly. Therefore, the structure of the covariance matrix of the received signal $\mathbf{y}(k)$ is

$$\mathrm{R}^{\mathbf{y}}_L = \mathcal{T}_L(\mathbf{H}_N)\,\mathrm{R}^a_{L+N-1}\mathcal{T}_L^H(\mathbf{H}_N) + \sigma_v^2 I_{mL} \tag{7}$$

0-7803-2431-5/95 $4.00 © 1995 IEEE

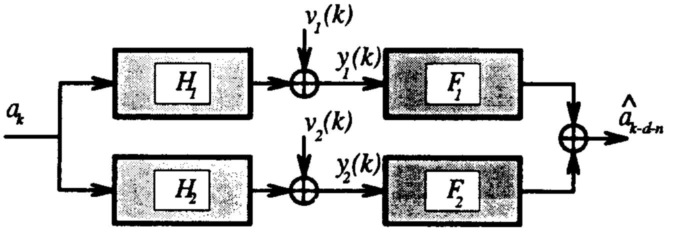

Figure 1: Polyphase representation of the single antenna T/m fractionally-spaced channel and equalizer or discrete-time representation of a symbol-rate sampled system with m antennas, for $m = 2$.

where $R^{\mathbf{Y}}_L = \mathrm{E}\mathbf{Y}_L(k)\mathbf{Y}^H_L(k)$ and $R^a_L = \mathrm{E}A_L(k)A^H_L(k)$. σ^2_v can be identified as the smallest eigenvalue of $R^{\mathbf{Y}}_L$. Replacing $R^{\mathbf{Y}}_L$ by $R^{\mathbf{Y}}_L - \sigma^2_v I_{mL}$ gives us the covariance matrix for noise-free data. Given the structure of $R^{\mathbf{Y}}_L$ in (7), the column space of $\mathcal{T}_L(\mathbf{H}_N)$ is called the signal subspace and its orthogonal complement the noise subspace [3].

2. MULTICHANNEL LINEAR PREDICTION (LP) AND EQUALIZATION

Consider first the noiseless case ($\sigma_v = 0$). And let the transmitted symbols be uncorrelated ($R^a_L = \sigma^2_a I_L$). When $L \geq \underline{L}$, $R^{\mathbf{Y}}_L$ is singular. If then L increases further by 1, the rank of $R^{\mathbf{Y}}_L$ increases by 1 and the dimension of its nullspace increases by $m-1$. Consider now the problem of (forward) predicting $\mathbf{y}(k)$ from $\mathbf{Y}_L(k-1)$ The prediction error can be written as

$$\begin{aligned}\tilde{\mathbf{y}}^f_L(k) &= \tilde{\mathbf{y}}(k)|_{\mathbf{Y}_{L(k-1)}} = \mathbf{P}_{\tilde{\mathbf{y}}^f_L}\mathbf{Y}_{L+1}(k) = \mathbf{y}(k) - \hat{\mathbf{y}}^f_L(k)\\ &= \mathbf{y}(k) - \hat{\mathbf{y}}(k)|_{\mathbf{Y}_{L(k-1)}} = [I_m \;\; -\mathbf{P}_{\hat{\mathbf{y}}^f_L}]\,\mathbf{Y}_{L+1}(k).\end{aligned} \qquad (8)$$

Minimizing the prediction error variance leads to the following optimization problem

$$\min_{\mathbf{P}_{\hat{\mathbf{y}}^f_L}} \left[I_m \;\; -\mathbf{P}_{\hat{\mathbf{y}}^f_L}\right] R^{\mathbf{Y}}_{L+1} \left[I_m \;\; -\mathbf{P}_{\hat{\mathbf{y}}^f_L}\right]^H = \sigma^2_{\tilde{\mathbf{y}}^f_L} \qquad (9)$$

or hence

$$\mathbf{P}_{\tilde{\mathbf{y}}^f_L} R^{\mathbf{Y}}_{L+1} = \left[\sigma^2_{\tilde{\mathbf{y}}^f_L} \;\; 0\cdots 0\right]. \qquad (10)$$

It can be shown that

$$\mathbf{P}_{\tilde{\mathbf{y}}^f_L}\mathcal{T}_{L+1}(\mathbf{H}_N) = \mathbf{h}(0)\;[1\;0\cdots 0]\;,\; \sigma^2_{\tilde{\mathbf{y}}^f_L} = \sigma^2_a\mathbf{h}(0)\mathbf{h}^H(0). \qquad (11)$$

In other words,

$$\frac{\mathbf{h}^H(0)}{\mathbf{h}^H(0)\mathbf{h}(0)}\mathbf{P}_{\tilde{\mathbf{y}}^f_L} = \mathbf{F}^{ZF}_{L+1,0}\,. \qquad (12)$$

3. 0 DELAY MMSE EQUALIZATION BY LP

When noise is present, MMSE equalization performs better than ZF equalization. For zero delay equalization, the following relation between the MMSE equalizer and linear prediction can readily be found:

$$\mathbf{F}^{MMSE}_{L+1,0} = \sigma^2_a\mathbf{h}^H(0)\sigma^{-2}_{\tilde{\mathbf{y}}^f_L}\mathbf{P}_{\tilde{\mathbf{y}}^f_L} \qquad (13)$$

The quantities $\mathbf{P}_{\tilde{\mathbf{y}}^f_L}$ and $\sigma^2_{\tilde{\mathbf{y}}^f_L}$ are readily determined from the linear prediction problem, while $\mathbf{h}(0)$ could be determined from the second-order statistics. Indeed, block subdiagonal $N-1$ of $R^{\mathbf{Y}}_{L+1}$ is given by $\sigma^2_a\mathbf{h}(N-1)\mathbf{h}^H(0)$. Alternatively, one could take the vector of prediction errors $\tilde{\mathbf{y}}(k)|_{\mathbf{Y}_{L(k-1)}}$ and determine an appropriate linear combination of these m signals by the Constant Modulus Algorithm.

4. MAXIMAL DELAY MMSE EQUALIZATION

The performance of the zero delay MMSE equalizer may be poor if the first channel coefficient $\mathbf{h}(0)$ is small. It is better to allow a delay so that the RHS of the normal equations that determine the MMSE equalizer contain all the channel coefficients. We propose a blind approach to obtain such an equalizer. One may remark that it would be possible to first estimate the channel coefficients, from which one can determine any equalizer. However, the emphasis here is on a simpler alternative. The approach consists of two steps:

1. do blind zero-delay ZF equalization. The equalizer output will be ($L+1 \geq \underline{L}$)

$$\hat{a}_k = a_k + \mathbf{F}^{ZF}_{L+1,0}\mathbf{V}_{L+1}(k)\,. \qquad (14)$$

2. obtain $\mathbf{F}^{MMSE}_{M,M-1}$ as a linear combination of a Wiener filter with $\mathbf{Y}_M(k)$ as input vector and $\hat{a}_{k-M+1}$ as desired response, and the backward linear prediction filter on the vector $\mathbf{Y}_M(k)$. Preferably, $M \geq N$.

For step 1, $\mathbf{F}^{ZF}_{L+1,0}$ can be found by LP in the noise-free case. In the case of additive (white) noise, the noise-free second-order statistics can easily be found after identifying the smallest eigenvalue of $R^{\mathbf{Y}}_{L+1}$. Alternatively, the Least-Squares approach of the noise-free case can be modified into a Total Least-Squares approach for the noisy case. Appropriate adaptive algorithms can be extrapolated from [6].Alternatively, the method discussed in the next section could be used for step 1.

For step 2, consider the FIR Wiener filtering problem

$$\min_{\mathbf{F}^W_M} \mathrm{E}\left|\hat{a}_{k-M+1} - \mathbf{F}^W_M\mathbf{Y}_M(k)\right|^2 \qquad (15)$$

which leads to the normal equations

$$\begin{aligned}\mathbf{F}^W_M R^{\mathbf{Y}}_M &= \mathrm{E}\hat{a}_{k-M+1}\mathbf{Y}^H_M(k)\\ &= \mathrm{E}a_{k-M+1}\mathbf{Y}^H_M(k) + \mathbf{F}^{ZF}_{L+1,0}\mathrm{E}\mathbf{V}_{L+1}(k-M+1)\mathbf{V}^H_M(k)\\ &= \mathbf{F}^{MMSE}_{M,M-1}R^{\mathbf{Y}}_M + [0\cdots 0\;\mathbf{f}(0)]\,\sigma^2_v\end{aligned} \qquad (16)$$

where $\mathbf{f}(0)$ represents the first m coefficients of $\mathbf{F}^{ZF}_{L+1,0}$. Consider now the multichannel backward prediction problem

$$\tilde{\mathbf{y}}^b_{M-1}(k) = \tilde{\mathbf{y}}(k-M+1)|_{\mathbf{Y}_{M-1}(k)} = \mathbf{P}_{\tilde{\mathbf{y}}^b_{M-1}}\mathbf{Y}_M(k) \qquad (17)$$

where $\mathbf{P}_{\widetilde{\mathbf{y}}^b_{M-1}} = [-\mathbf{P}_{\widehat{\mathbf{y}}^b_{M-1}} \; I_m]$, with normal equations

$$\mathbf{P}_{\widetilde{\mathbf{y}}^b_{M-1}} \mathrm{R}^{\mathbf{y}}_M = \left[0 \cdots 0 \; \sigma^2_{\widetilde{\mathbf{y}}^b_{M-1}}\right] . \quad (18)$$

From (16) and (18), we conclude

$$\mathbf{F}^{MMSE}_{M,M-1} = \mathbf{F}^W_M - \sigma^2_v \mathbf{f}(0) \sigma^{-2}_{\widetilde{\mathbf{y}}^b_{M-1}} \mathbf{P}_{\widetilde{\mathbf{y}}^b_{M-1}} . \quad (19)$$

In this expression for $\mathbf{F}^{MMSE}_{M,M-1}$, all quantities can easily be found by adaptive filtering except perhaps σ^2_v. But an estimate for σ^2_v results as a byproduct of step 1.

5. CONSTRAINED IIR FILTER DFE

Whereas the equalizers considered in the previous sections are linear, here we consider an equalizer structure with decision feedback. The approach is in fact a multichannel extension of the adaptive notch filter approach for sinusoids in noise. As a consequence, the method will continue to work well even if the additive noise and/or the transmitted symbols are colored.

Let $P_{\widehat{\mathbf{y}}^f_L}(z)$ and $P_{\widetilde{\mathbf{y}}^f_L}(z)$ be the z-transforms of the forward prediction and prediction error filters (of the noise-free case) so that $P_{\widetilde{\mathbf{y}}^f_L}(z) = I_M - z^{-1} P_{\widehat{\mathbf{y}}^f_L}(z)$. To alleviate the notation, $P(z)$ will continue to represent $P_{\widetilde{\mathbf{y}}^f_L}(z)$ and let q^{-1} be the unit delay operator: $q^{-1}\mathbf{y}(k) = \mathbf{y}(k-1)$. Since $P(z)H(z) = \mathbf{h}(0)$, the noise-free received vector signal $\mathbf{y}(k) = H(q)a_k$, which is a multichannel MA process, is also a (singular) multichannel AR process: $P(q)\mathbf{y}(k) = \mathbf{h}(0)a_k$. For the noisy received signal $\mathbf{y}(k) = H(q)a_k + \mathbf{v}(k)$, we get

$$P(q)\mathbf{y}(k) = \mathbf{h}(0)a_k + P(q)\mathbf{v}(k) \quad (20)$$

which is a constrained multichannel ARMA process, apart from the term $\mathbf{h}(0)a_k$. In the scalar case, the prediction error filter is minimum-phase. For the multichannel case, the extension is that $\det[P(z)]$ is minimum-phase. However since in the noise-free case $\mathrm{R}^{\mathbf{y}}_{L+1}$ is singular, zeros on the unit circle can occur. Hence, if we want to use (20) to recover $\mathbf{v}(k)$, we need to introduce a damping factor $\rho \lessapprox 1$:

$$\mathbf{v}(k) = P^{-1}(q/\rho)[P(q)\mathbf{y}(k) - \mathbf{h}(0)a_k] . \quad (21)$$

It would perhaps be preferable to replace $P^{-1}(q/\rho)$ by $\mathrm{Adj}[P(q)]/\det[P(q/\rho)]$, but (21) can be more straightforwardly implemented by the following procedure

$$\begin{cases} \mathbf{s}(k) = P_{\widetilde{\mathbf{y}}^f_L}(q)\mathbf{y}(k) - \rho P_{\widehat{\mathbf{y}}^f_L}(q/\rho)\widehat{\mathbf{v}}(k-1) \\ \widehat{a}_k = \mathrm{dec}[\frac{\mathbf{h}^H(0)}{\mathbf{h}^H(0)\mathbf{h}(0)}\mathbf{s}(k)] \\ \widehat{\mathbf{v}}(k) = \mathbf{s}(k) - \mathbf{h}(0)\widehat{a}_k \end{cases} \quad (22)$$

where dec denotes the decision operation, whose argument is ideally $a_k + \frac{\mathbf{h}^H(0)}{\mathbf{h}^H(0)\mathbf{h}(0)}\mathbf{v}(k)$. Various algorithms are now possible to adapt the coefficients $\mathbf{P}_{\widehat{\mathbf{y}}^f_L}$ such as the Recursive Prediction Error Method and its simplifications.

6. n-STEP AHEAD LINEAR PREDICION

In section 2, we have shown how $\mathbf{F}^{ZF}_{L+1,0}$ can be obtained by linear prediction (LP) in the noise-free case. In the noisy case, but with high SNR, one might be tempted to continue to pursue the LP approach by e.g. replacing $\mathbf{h}(0)$ in (12) by the eigenvector of $\sigma^2_{\widetilde{\mathbf{y}}^f_L}$ corresponding to its largest eigenvalue. This approach will clearly not work very well if $\mathbf{h}(0)$ is small, which can occur since the channel is not necessarily minimum-phase. In the approach outline below, a ZFE with arbitrary delay is obtained in the noise-free case, involving a corresponding channel coefficient.

Consider n-step ahead (forward) linear prediction of the noise-free $\mathbf{y}(k)$ of order $L \geq \underline{L}$:

$$\widetilde{\mathbf{y}}^f_{L,n}(k) = \widetilde{\mathbf{y}}(k)|_{\mathbf{Y}_{L(k-n)}} = P_{\widetilde{\mathbf{y}}^f_{L,n}}(q)\mathbf{y}(k) . \quad (23)$$

The use of the optimal predictor will result in

$$\widetilde{\mathbf{y}}^f_{L,n}(k) = P_{\widetilde{\mathbf{y}}^f_{L,n}}(q)\mathbf{y}(k) = \sum_{i=0}^{n-1} \mathbf{h}(i)\, a_{k-i} . \quad (24)$$

For $n = 1$, we find the results of section 2. Note that we can regard $\widetilde{\mathbf{y}}^f_{L,n}(k)$ as the received signal from a truncated channel. If we now apply backward linear prediction of sufficient order M (replace N by n in the expression in (5) for $\underline{L}$) to the signal $\widetilde{\mathbf{y}}^f_{L,n}(k)$, then we obtain as optimal prediction error

$$\widetilde{\widetilde{\mathbf{y}}}^b_M(k) = P_{\widetilde{\widetilde{\mathbf{y}}}^b_M}(q)\widetilde{\mathbf{y}}^f_{L,n}(k) = \mathbf{h}(n-1)a_{k-n+1} . \quad (25)$$

From (25) and (23), we deduce that we can obtain a ZF equalizer with delay $n-1$ as

$$F^{ZF}_{n+L+M,n-1}(q) = \frac{\mathbf{h}^H(n-1)}{\mathbf{h}^H(n-1)\mathbf{h}(n-1)} P_{\widetilde{\widetilde{\mathbf{y}}}^b_M}(q) P_{\widetilde{\mathbf{y}}^f_{L,n}}(q) . \quad (26)$$

It is true that any $\mathbf{h}(n-1)$ could be small. So the way these results should be used is perhaps through the combination of several ZFEs corresponding to several delays ($= 0, 1, \ldots, n$), such that the chance for $\sum_{i=0}^{n} \|\mathbf{h}(i)\|^2$ being small becomes small. The outputs of these ZFEs should be properly delayed to align them at the same a_{k-n}.

7. DML CHANNEL ESTIMATION

Consider additive white Gaussian noise but deterministic (D) transmitted symbols a_k. Maximizing the likelihood function for the data $\mathbf{Y}_M(k)$ leads to the following separable LS problem

$$\min_{\mathbf{H}_N, A_{M+N-1}(k)} \|\mathbf{Y}_M(k) - \mathcal{T}_M(\mathbf{H}_N)\, A_{M+N-1}(k)\|^2_2 . \quad (27)$$

Eliminating $A_{M+N-1}(k)$ in terms of $\mathbf{H}_N$, we get

$$\min_{\mathbf{H}_N} \|P^{\perp}_{\mathcal{T}_M(\mathbf{H}_N)}\mathbf{Y}_M(k)\|^2_2 . \quad (28)$$

In order to find an attractive iterative procedure for solving this optimization problem, we should work with a minimal parameterization of the noise subspace, which we have obtained in [4] from the prediction problem of the noise-free signal. We get $P^{\perp}_{\mathcal{T}_M(\mathbf{H}_N)} = P_{\mathcal{G}^H_M(G_N)}$. The number of degrees of freedom in $\mathbf{H}_N$ and G_N is both $mN-1$ (the proper scaling factor cannot be determined). So $\mathbf{H}_N$ can be uniquely determined from G_N and vice versa. Hence, we can reformulate the optimization problem in (28) as

$$\min_{G_N} \|P_{\mathcal{G}^H_M(G_N)} \mathbf{Y}_M(k)\|_2^2 \ . \quad (29)$$

Due to the (almost) block Toeplitz character of $\mathcal{G}_M$, the product $\mathcal{G}_M \mathbf{Y}_M(k)$ represents a convolution. Due to the commutativity of convolution, we can write $\mathcal{G}_M(G_N)\mathbf{Y}_M(k) = \mathcal{Y}_N(\mathbf{Y}_M(k))[1 \ \ G_N^H]^H$ for some properly structured $\mathcal{Y}_N(\mathbf{Y}_M(k))$. This leads us to rewrite (29) as the following IQML problem

$$\min_{G_N} \begin{bmatrix} 1 \\ G_N \end{bmatrix}^H \mathcal{Y}_N^H(\mathbf{Y}_M(k)) \left(\mathcal{G}_M(G_N)\mathcal{G}_M^H(G_N)\right)^{-1} \mathcal{Y}_N(\mathbf{Y}_M(k)) \begin{bmatrix} 1 \\ G_N \end{bmatrix} \quad (30)$$

An initial estimate may be obtained from a subspace fitting approach [3] or from the LP problem. Such an initial estimate is consistent and hence one iteration of (30) will be sufficient to generate an estimate that is asymptotically equivalent to the global optimizer of (30). Cramer-Rao bounds have been obtained and analyzed in [3].

The characterization $\mathcal{G}_M(G_N)$ of the noise subspace is not very robust because N (and $h_1(N-1) \neq 0$) is assumed known. Note however that asymptotically the DML criterion (29) becomes the sum of the squares of $\mathbf{w}(k)$ in

$$\overline{P}(q)\mathbf{y}(k) = \overline{P}(q)\mathbf{v}(k) = \left(\overline{P}(q)\overline{P}^H(1/q^*)\right)^{1/2} \mathbf{w}(k) \quad (31)$$

obtained from (20), where $\overline{P}(q) = \mathbf{h}^{\perp H}(0)P(q)$, $\mathbf{h}^{\perp}(0)$ is a $m \times (m-1)$ matrix of rank $m-1$ s.t. $\mathbf{h}^{\perp H}(0)\mathbf{h}(0) = 0$, and $(.)^{1/2}$ is a minimum-phase factor of its argument. This leads us to introduce a more robust approximate DML as

$$\min_{\overline{\mathbf{P}}} \|P_{\mathcal{T}^H_{M-L}(\overline{\mathbf{P}})} \mathbf{Y}_M(k)\|_2^2 \quad (32)$$

for any $L \geq \underline{L}$. (32) can again be solved in the IQML fashion as in (30). A minimal parameterization for $\overline{\mathbf{P}}$ is $\mathbf{h}^{\perp H}(0)\mathbf{P} = [\,[I_{m-1}\ g]\mathcal{P} \ \ Q]$ where g $((m-1) \times 1)$ and Q $((m-1) \times mL)$ are the free parameters and $\mathcal{P}$ is a permutation matrix that permutes g into the column of $\mathbf{h}^{\perp H}(0)$ that corresponds to the largest element of (an LP based estimate of) $\mathbf{h}(0)$.

8. GML: ML WITH GAUSSIAN PRIOR

Computer simulations have shown that for small m, the channel estimate from DML can be relatively bad if the channel impulse response tapers off near the ends. As an intermediate step to the computationally expensive stochastic ML in which the discrete distribution of the a_k is exploited, we propose to introduce a Gaussian i.i.d. prior for the a_k. The GML criterion is

$$\min_{\mathbf{H}_N, A_{M+N-1}(k)} \left\| \begin{bmatrix} \mathbf{Y}_M(k) \\ 0_{N+M-1} \end{bmatrix} - \begin{bmatrix} \mathcal{T}_M(\mathbf{H}_N) \\ \frac{\sigma_v}{\sigma_a} I_{N+M-1} \end{bmatrix} A_{M+N-1}(k) \right\|_2^2 \quad (33)$$

The Cramer-Rao bound for the channel $CRB_{\widehat{\mathbf{H}}_N^{tT}}$ for the problem (33) can be found to be:

$$C(\widehat{\mathbf{H}}_N^{tT}) = \sigma_v^2 \left[\mathcal{K}^H_{M,N}(k) P^{\perp}_{\Phi_M(\mathbf{H}_N)} \mathcal{K}_{M,N}(k)\right]^{-1} , \quad (34)$$

$\Phi_M(\mathbf{H}_N) = \begin{bmatrix} \mathcal{T}_M(\mathbf{H}_N) \\ \frac{\sigma_v}{\sigma_a} I_{N+M-1} \end{bmatrix}$, $\mathcal{K}_{M,N}(k) = \begin{bmatrix} \mathcal{A}_{M,N}(k) \\ 0_{N+M-1,mN} \end{bmatrix}$, $\mathcal{A}_{M,N}(k) = A_{M,N}(k) \otimes I_m$ and

$$A_{M,N}(k) = \begin{bmatrix} a(k) & \cdots & a(k-N+1) \\ \vdots & \iddots & \vdots \\ a(k-M+1) & \cdots & a(k-M-N+2) \end{bmatrix} . \quad (35)$$

Due to the Gaussian prior, the singularity in the CRB gets eliminated w.r.t. the DML problem. Moreover, as computer simulations have shown, the CRB for the GML problem remains small even for channels whose impulse response tapers off near the ends. A simple way to find the channel estimate from the GDML problem is to start with the DML problem. In the DML problem, the estimate of the data corresponds to a ZFE, while in the GML problem, the data estimate corresponds to a MMSE equalizer. Using the estimates for the channel and for σ_v/σ_a from DML, apply the MMSE equalizer and then estimate the channel again with these estimated data from the LS problem in 27). Since the DML method leads to a consistent channel estimate, this one iteration will lead to a channel estimate that can asymptotically not be distinguished from the GDML estimate.

9. 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] D.T.M. Slock. "Blind Fractionally-Spaced Equalization, Perfect-Reconstruction Filter Banks and Multichannel Linear Prediction". In *Proc. ICASSP 94 Conf.*, Adelaide, Australia, April 1994.

[3] D.T.M. Slock and C.B. Papadias. "Blind Fractionally-Spaced Equalization Based on Cyclostationarity". In *Proc. Vehicular Technology Conf.*, Stockholm, Sweden, June 1994.

[4] D.T.M. Slock. "Subspace Techniques in Blind Mobile Radio Channel Identification and Equalization using Fractional Spacing and/or Multiple Antennas". In *Proc. 3rd International Workshop on SVD and Signal Processing*, Leuven, Belgium, Aug. 22-25 1994.

[5] D.T.M. Slock. "Blind Joint Equalization of Multiple Synchronous Mobile Users Using Oversampling and/or Multiple Antennas". In *Proc. 28th Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, Oct. 31 - Nov. 2 1994.

[6] C.E. Davila. "An Efficient Recursive Total Least Squares Algorithm for FIR Adaptive Filtering". *IEEE Trans. Signal Processing*, SP-42(2):268–280, Feb. 1994.