MULTI-STEP LINEAR PREDICTORS-BASED BLIND EQUALIZATION OF MULTIPLE-INPUT MULTIPLE-OUTPUT CHANNELS

Jitendra K. Tugnait *Bin Huang*

Department of Electrical Engineering
Auburn University, Auburn, Alabama 36849, USA
E-mail: tugnait@eng.auburn.edu huangbi@eng.auburn.edu

ABSTRACT

Blind equalization of MIMO (multiple-input multiple-output) communications channels is considered using primarily the second-order statistics of the data. In several applications the underlying equivalent discrete-time mathematical model is that of a MIMO linear system where the number of inputs equals the number of users (sources) and the number of outputs is related to the number of sensors and the sampling rate. Recently we investigated the structure of multi-step linear predictors for IIR/FIR MIMO systems with irreducible transfer functions and derived an upper bound on its length (Tugnait, 1998 IEEE DSP Workshop). In past multi-step linear predictors have been considered in the literature only for single-input multiple-output models. In this paper we apply the results of (Tugnait, 1998 IEEE DSP Workshop) for blind equalization of MIMO channels using MMSE linear equalizers. Extensions to the case where the "subchannel" transfer functions have common zeros/factors is also investigated. An illustrative simulation example is provided

1. INTRODUCTION

Consider a discrete-time IIR MIMO system with N outputs and M inputs:

$$\mathbf{y}(k) \;=\; \mathcal{F}(z)\mathbf{w}(k) + \mathbf{n}(k) \;=\; \mathbf{s}(k) + \mathbf{n}(k) \qquad (1)$$

where $\mathbf{y}(k) = [y_1(k) \vdots y_2(k) \vdots \cdots \vdots y_N(k)]^T$, similarly for $\mathbf{w}(k)$, $\mathbf{s}(k)$ and $\mathbf{n}(k)$, z is the $\mathcal{Z}-$transform variable as well as the backward-shift operator (i.e., $z^{-1}\mathbf{w}(k) = \mathbf{w}(k-1)$, etc.), $\mathbf{s}(k)$ is the noise-free output, $\mathbf{n}(k)$ is the additive measurement noise and the $N \times M$ matrix $\mathcal{F}(z)$ is given by

$$\mathcal{F}(z) \;=\; \mathcal{A}^{-1}(z)\mathcal{B}(z),$$

$$\mathcal{A}(z) = I + \sum_{i=1}^{n_a} \mathbf{A}_i z^{-i} \;\text{ and }\; \mathcal{B}(z) \;=\; \sum_{i=0}^{n_b} \mathbf{B}_i z^{-i}. \qquad (2)$$

We allow all of the above variables to be complex-valued. The following assumptions are made regarding (1)-(2):

- **(H1)** $N > M$.
- **(H2)** Rank$\{\mathcal{B}(z)\} = M$ $\forall z$ including $z = \infty$ but excluding $z = 0$, i.e., $\mathcal{B}(z)$ is irreducible [5, Sec. 6.3].
- **(H3)** Input sequence $\{\mathbf{w}(k)\}$ is zero-mean, spatially independent and temporally i.i.d. with each of its components having non-zero fourth cumulants. Take $E\{\mathbf{w}(k)\mathbf{w}^H(k)\} = I_M$ where I_M is the $M \times M$ identity matrix and the superscript H is the Hermitian operation.
- **(H4)** $\{\mathbf{n}(k)\}$ is zero-mean Gaussian with $E\{\mathbf{n}(k+\tau)\mathbf{n}^H(k)\} = \sigma_n^2 I_N \delta(\tau)$.
- **(H5)** $\det(\mathcal{A}(z)) \neq 0$ for $|z| \geq 1$.

This work was supported by ONR Grant N00014-97-1-0822 and by NSF Grant CCR-9803850.

Notations and Definitions: Let $\mathcal{B}^{(l)}(z)$ denote the $l-$th column of $\mathcal{B}(z)$ such that $\mathcal{B}^{(l)}(z) = \sum_{i=0}^{L_l} \mathbf{B}_i^{(l)} z^{-i}$ where $L_l = \deg\left(\mathcal{B}^{(l)}(z)\right) =$ lowest degree of the polynomial $\mathcal{B}^{(l)}(z)$. By (2), $L_l \leq n_b$ $\forall l$. The polynomial matrix $\mathcal{B}(z)$ is said to be column-reduced if rank $\left\{[\mathbf{B}_{L_1}^{(1)} \vdots \cdots \vdots \mathbf{B}_{L_M}^{(M)}]\right\} = M$ [5]. Consider the Hilbert space $\mathcal{H}$ of square integrable complex random variables on a common probability space endowed with the inner product (for scaler complex random variables x_1 and x_2) $< x_1, x_2 >= E\{x_1 x_2^*\}$ where the superscript $*$ denotes complex conjugation (see [4]). Let $Sp\{x_i \in I\}$ denote the subspace of $\mathcal{H}$ generated by the random variables/vectors in the set $\{x_i \in I\}$. Given an $N-$variate $\mathbf{s}(k)$ with $i-$th component $s_i(k)$, define the subspace

$$H_{k-m;L_1,L_2,\cdots,L_N}(\mathbf{s}) :=$$
$$Sp\{\; s_i(k-l_i),\;\; m \leq l_i \leq L_i;\;\; i = 1, 2, \cdots, N\}.$$

We will use $H_{k-m}(\mathbf{s})$ to denote $H_{k-m;\infty,\cdots,\infty}(\mathbf{s})$. Let $(\mathbf{s}(k)|H_{k-1}(\mathbf{s}))$ denote the orthogonal projection of $\mathbf{s}(k)$ onto the subspace $H_{k-1}(\mathbf{s})$ [4]. $\square$

Models such as (1)-(2) with $\mathcal{F}(z) = \mathcal{B}(z)$ arise in several useful digital communications and other applications [1]-[3], [6],[7] where one of the objectives is to estimate the multichannel impulse response $\{\mathbf{B}_i\}$ and/or to recover the inputs $\mathbf{w}(k)$ given the noisy measurements but not given the knowledge of the system transfer function. One of the popular approaches is that using linear prediction [1]-[3] where existence of finite-length one-step linear predictors plays a key role. In the MIMO case it is known that under (**H1**)-(**H3**), finite-length one-step linear predictors exist for the process $\mathbf{s}(k)$ [6]. The length (or an upper-bound on it) has not been specified in [6]. Under an additional condition that $\mathcal{B}(z)$ is column-reduced, it is stated in [1] that there exists a linear predictor (for $\mathbf{s}(k)$) of length no longer than $\sum_{i=1}^{M} L_i$.

A one-step linear prediction-based approach was first proposed in [9] and later expanded upon in [2]. Unlike the subspace-based methods of [10], [11] and others, the linear prediction (LP) based approach of [9] and [2] turns out to be rather insensitive to the order of the underlying FIR channel (so long as one overfits). More recently, it has been pointed out in [12] and [13] that the LP-based approach can be further significantly improved by utilizing some additional information not exploited by LP. In this paper we will follow the approach of [13] which is based upon multi-step linear prediction. Unlike [13] we allow multiple inputs and IIR channels. Unlike [12] we allow MIMO transfer functions that are not column-reduced and we also allow IIR channels.

2. BLIND EQUALIZATION

2.1. Finite-Length Multi-Step Predictors

Here we first summarize the results of [14] in Theorem 1. Let

$$\mathcal{F}(z) \;=\; \sum_{i=0}^{\infty} \mathbf{F}_i z^{-i}. \qquad (3)$$

It has been shown in [14] that using $\mathbf{s}(k) = \mathcal{F}(z)\mathbf{w}(k)$, we obtain for appropriate choices of $\mathbf{A}_{li}$'s and $\mathbf{B}_{li}$'s ($l \geq 1$)

$$\mathbf{s}(k) = -\sum_{i=l}^{n_a+l-1} \mathbf{A}_{li}\mathbf{s}(k-i) + \sum_{i=0}^{n_b+l-1} \mathbf{B}_{li}\mathbf{w}(k-i) \quad (4)$$

such that

$$\mathbf{B}_{li} \;=\; \mathbf{F}_i \quad \text{for} \quad 0 \leq i \leq l-1. \quad (5)$$

Let us rewrite (4) as

$$\mathbf{s}(k) = \mathbf{e}(k|k-l) + \widehat{\mathbf{s}}(k|k-l), \quad (6)$$

$$\mathbf{e}(k|k-l) := \sum_{i=0}^{l-1} \mathbf{B}_{li}\mathbf{w}(k-i) = \sum_{i=0}^{l-1} \mathbf{F}_i\mathbf{w}(k-i), \quad (7)$$

$$\widehat{\mathbf{s}}(k|k-l) := -\sum_{i=l}^{n_a+l-1} \mathbf{A}_{li}\mathbf{s}(k-i) + \sum_{i=l}^{n_b+l-1} \mathbf{B}_{li}\mathbf{w}(k-i). \quad (8)$$

$\boxed{\textbf{Theorem 1}}$ [14]. Under (**H1**)-(**H3**), (**H5**), and for $l = 1, 2, \cdots$, $\{\mathbf{s}(k)\}$ can be decomposed as in (6) such that

$$E\{\mathbf{e}(k|k-l)\mathbf{s}^H(k-m)\} \;=\; 0 \quad \forall m \geq l, \quad (9)$$

$$\widehat{\mathbf{s}}(k|k-l) \;=\; (\mathbf{s}(k)|H_{k-l}(\mathbf{s}))\,, \quad (10)$$

$$\widehat{\mathbf{s}}(k|k-l) \in H_{k-l;n_a+Mn_b+l-1,\cdots,n_a+Mn_b+l-1}(\mathbf{s}) \quad (11)$$

and

$$\widehat{\mathbf{s}}(k|k-l)$$
$$= (\mathbf{s}(k)|H_{k-l;n_a+Mn_b+l-1,\cdots,n_a+Mn_b+l-1}(\mathbf{s})) \quad \bullet \quad (12)$$

It follows from Theorem 1 that

$$\widehat{\mathbf{s}}(k|k-l) = \sum_{i=l}^{P_l} \overline{\mathbf{A}}_{li}\mathbf{s}(k-i) \;\text{ where }\; P_l \geq n_a + Mn_b + l - 1, \quad (13)$$

for some $N \times N$ matrices $\overline{\mathbf{A}}_{li}$s. By (6) and (9) (recall also the orthogonal projection theorem), we have

$$\widehat{\mathbf{s}}(k|k-l) \;=\; \arg\left\{\min_{\mathbf{x}(k)\in H_{k-l}(\mathbf{s})} E\{\|\mathbf{s}(k)-\mathbf{x}(k)\|^2\right\}. \quad (14)$$

Therefore, $\widehat{\mathbf{s}}(k|k-l)$ is the $l-$step (ahead) linear predictor of $\mathbf{s}(k)$ given $\{\mathbf{s}(m),\;\; m \leq k-l\}$. By (12) it is also the $l-$step (ahead) linear predictor of $\mathbf{s}(k)$ given $\{\mathbf{s}(m),\;\; k-P_l \leq m \leq k-l\}$. Using (6) and (13) we have

$$\mathbf{s}(k) \;=\; \sum_{i=l}^{P_l} \mathbf{A}_i^{(l)}\mathbf{s}(k-i) \;+\; \mathbf{e}(k|k-l). \quad (15)$$

By (9) and (15), for $m \geq l$,

$$E\{\mathbf{s}(k)\mathbf{s}^H(k-m)\} \;=\; \sum_{i=l}^{P_l} \mathbf{A}_i^{(l)} E\{\mathbf{s}(k-i)\mathbf{s}^H(k-m)\}. \quad (16)$$

By the orthogonal projection theorem and (12), it is sufficient to consider (16) for $m = l, l+1, \cdots, P_l$ in order to solve for $\mathbf{A}_i^{(l)}$s. Using these values of m in (16) we may write

$$\left[\begin{array}{ccc} \mathbf{A}_l^{(l)} & \cdots & \mathbf{A}_{P_l}^{(l)} \end{array}\right] \mathcal{R}_{ss(P_l-l)} = \left[\begin{array}{ccc} \mathbf{R}_{ss}(l) & \cdots & \mathbf{R}_{ss}(P_l) \end{array}\right] \quad (17)$$

where $\mathcal{R}_{ssL}$ denotes a $[N(L+1)] \times [N(L+1)]$ matrix with its ij-th block element as $\mathbf{R}_{ss}(j-i) = E\{\mathbf{s}(k+j-i)\mathbf{s}^H(k)\}$.

Note that $\mathcal{R}_{ss(P_l-l)}$ is not necessarily full rank, therefore, the coefficients $\mathbf{A}_i^{(l)}$s are not necessarily unique. A minimum norm solution to (17) may be obtained as

$$\left[\begin{array}{ccc} \mathbf{A}_l^{(l)} & \cdots & \mathbf{A}_{P_l}^{(l)} \end{array}\right] = \left[\begin{array}{ccc} \mathbf{R}_{ss}(l) & \cdots & \mathbf{R}_{ss}(P_l) \end{array}\right] \mathcal{R}^{\#}_{ss(P_l-l)} \quad (18)$$

where the superscript $\#$ denotes the pseudoinverse.

2.2. Estimation of Noise Variance

In a fashion similar to $\mathcal{R}_{ssL}$ in (17), let $\mathcal{R}_{yyL}$ denote a $[N(L+1)] \times [N(L+1)]$ matrix with its ij-th block element as $\mathbf{R}_{yy}(j-i) = E\{\mathbf{y}(k+j-i)\mathbf{y}^H(k)\}$; define similarly $\mathcal{R}_{nnL}$ pertaining to the additive noise. Carry out an eigendecomposition of $\mathcal{R}_{yyP_1}$. Then the smallest $N-M$ eigenvalues of $\mathcal{R}_{yyP_1}$ equal σ_n^2 [14]. Therefore, a consistent estimate $\widehat{\sigma}_n^2$ of σ_n^2 is obtained by taking it as the average of the smallest $N-M$ eigenvalues of $\widehat{\mathcal{R}}_{yyP_1}$, the data-based consistent estimate of $\mathcal{R}_{yyP_1}$. We will need the estimate of noise variance later to calculate $\mathcal{R}^{\#}_{ss(P_l-l)}$ in (18) for various $l \geq 1$. By (13), $P_l - l \geq n_a + Mn_b - 1$, independent of l. This suggests that we keep

$$P_l - l \;=\; \overline{L} \;\geq\; n_a + Mn_b - 1 \quad (\forall\, l). \quad (19)$$

Then under (**H3**) and (**H4**),

$$\mathcal{R}_{ss\overline{L}} \;=\; \mathcal{R}_{yy\overline{L}} - \mathcal{R}_{nn\overline{L}} \;=\; \mathcal{R}_{yy\overline{L}} \;-\; \sigma_n^2 I_{N(\overline{L}+1)}. \quad (20)$$

Thus, $\mathcal{R}_{ss\overline{L}}$ can be estimated from noisy data.

2.3. MMSE Equalizer with Delay d

We wish to design an MMSE (minimum mean-square error) linear equalizer of a specified length. Using the orthogonality principle [4], it follows that the MMSE equalizer of length L_e+1 to estimate $\mathbf{w}(k-d)$ ($d \geq 0$) based upon $\mathbf{y}(n),\;\; n = k, k-1, \cdots, k-L_e$, satisfies

$$\left[\begin{array}{cccc} \overline{\mathbf{G}}_{d,0} & \overline{\mathbf{G}}_{d,1} & \cdots & \overline{\mathbf{G}}_{d,L_e} \end{array}\right] \;=$$
$$\left[\begin{array}{cccccc} \mathbf{F}_d^H & \mathbf{F}_{d-1}^H & \cdots & \mathbf{F}_0^H & 0 & \cdots & 0 \end{array}\right] \mathcal{R}^{-1}_{yyL_e} \quad (21)$$

The equalized output is given by

$$\widehat{\mathbf{w}}(k-d) \;=\; \sum_{i=0}^{L_e} \overline{\mathbf{G}}_{d,i}\mathbf{y}(k-i). \quad (22)$$

Clearly one can obtain a consistent estimate of $\mathcal{R}_{yyL_e}$ from the given data. It remains to estimate $\mathbf{F}_l$'s to complete the design. This is where the multistep predictor approach turns out to be useful.

2.4. Partial Channel Identification

As noted in Sec. 2.3, we need estimates of $\mathbf{F}_i$ for $i = 0, 1, \cdots, d$. We now show how Sec. 2.1 may be exploited for this purpose. Let

$$\widetilde{L} \;=\; \overline{L} + d + 1 \quad (23)$$

where $\overline{L}$ is as in (19). Rewrite (15) as

$$\mathbf{e}(k|k-l) \;=\; \sum_{i=0}^{\widetilde{L}} \overline{\mathbf{A}}_i^{(l)}\mathbf{s}(k-i) \quad (24)$$

where

$$\overline{\mathbf{A}}_i^{(l)} \;=\; \begin{cases} I_N & \text{for} \quad i = 0 \\ 0 & \text{for} \quad 1 \leq i \leq l-1 \\ -\mathbf{A}_i^{(l)} & \text{for} \quad l \leq i \leq \overline{L}+l \\ 0 & \text{for} \quad \overline{L}+l+1 \leq i \leq \widetilde{L}\,. \end{cases} \quad (25)$$

By (19) $P_l = \overline{L}+l$, therefore, for each l, we estimate $\overline{L}+1$ coefficients in (18). For $l \geq 2$, define

$$\overline{\mathbf{e}}_l(k) := \mathbf{e}(k|k-l) - \mathbf{e}(k|k-l+1) = \sum_{i=0}^{\widetilde{L}} \mathbf{D}_i^{(l,l-1)}\mathbf{s}(k-i) \quad (26)$$

where

$$\mathbf{D}_i^{(l,l-1)} := \overline{\mathbf{A}}_i^{(l)} - \overline{\mathbf{A}}_i^{(l-1)}, \quad i = 0,1,\cdots,\widetilde{L}. \quad (27)$$

By (25), $\mathbf{D}_0^{(l,l-1)} = 0 \ \forall l \geq 2$.

Consider the $[N(d+1)]$−vector

$$\mathbf{E}(k) := \left[\overline{\mathbf{e}}_{d+1}^T(k+d) \vdots \cdots \vdots \overline{\mathbf{e}}_2^T(k+1) \vdots \mathbf{e}^T(k|k-1) \right]^T. \quad (28)$$

Using (24)-(28) we have

$$\mathbf{E}(k) = \mathcal{D}S(k) \quad (29)$$

where

$$S(k) := \left[\mathbf{s}^T(k+d-1) \vdots \mathbf{s}^T(k+d-2) \vdots \cdots \vdots \mathbf{s}^T(k-\widetilde{L}) \right]^T \quad (30)$$

is a $[N(\widetilde{L}+d)]$−column vector and $\mathcal{D}$ is a $[N(d+1)] \times [N(\widetilde{L}+d)]$ matrix composed of $\mathbf{D}_i^{(\cdot,\cdot)}$s and $\mathbf{A}_i^{(1)}$s. Using (7), (26) and (28), we have

$$\mathbf{E}(k) = \begin{bmatrix} \mathbf{F}_d^T & \mathbf{F}_{d-1}^T & \cdots & \mathbf{F}_0^T \end{bmatrix}^T \mathbf{w}(k) =: \widetilde{\mathbf{F}}\mathbf{w}(k). \quad (31)$$

By (29), (31) and (**H3**), it follows that

$$\mathbf{R}_{EE}(0) = E\{\mathbf{E}(k)\mathbf{E}^H(k)\} = \widetilde{\mathbf{F}}\widetilde{\mathbf{F}}^H = \mathcal{D}\mathcal{R}_{ss(\widetilde{L}+d)}\mathcal{D}^H \quad (32)$$

Since $\widetilde{\mathbf{F}}\widetilde{\mathbf{F}}^H = \widetilde{\mathbf{F}}\mathbf{U}\mathbf{U}^H\widetilde{\mathbf{F}}^H$ for any unitary $\mathbf{U}$, one can not uniquely determine $\widetilde{\mathbf{F}}$ from (32); one needs to exploit the higher-order statistics of the data [8]. Clearly rank$(\widetilde{\mathbf{F}}\widetilde{\mathbf{F}}^H) = M$. Calculate $\mathbf{R}_{EE}(0)$ as

$$\mathbf{R}_{EE}(0) = \mathcal{D}\left[\mathcal{R}_{yy(\widetilde{L}+d-1)} - \sigma_n^2 I_{N(\widetilde{L}+d)}\right]\mathcal{D}^H. \quad (33)$$

Carry out an eigendecomposition of $\mathbf{R}_{EE}(0)$. Let $\mathbf{l}_i$ ($i = 1,2,\cdots,M$) denote the orthonormal eigenvectors of $\mathbf{R}_{EE}(0)$ corresponding to the M largest eigenvalues γ_i's. Set $\mathcal{L} = \left[\mathbf{l}_1 \vdots \cdots \vdots \mathbf{l}_M\right]$ and , $= diag\left(\sqrt{\gamma_1},\cdots,\sqrt{\gamma_M}\right)$. Set $\mathbf{H} = \mathcal{L}$, . Then $\mathbf{H}\mathbf{H}^H = \widetilde{\mathbf{F}}\widetilde{\mathbf{F}}^H$. Define $\mathbf{z}(k) = \mathbf{H}^\#\mathbf{E}(k)$ where $\mathbf{H}^\# =$, $^{-1}\mathcal{L}^{\mathcal{H}}$. Apply the joint diagonalization procedure of [8] to estimate the unitary matrix $\mathbf{U}$ exploiting the fourth-order cumulants of $\mathbf{z}(k)$ (at zero lag). Then

$$\mathbf{F} = \mathbf{H}\mathbf{U}. \quad (34)$$

2.5. Practical Implementation

Given data $\mathbf{y}(k)$, $k = 1,2,\cdots,T$. Pick the length L_e+1 and delay d of the MMSE equalizer. (By Theorem 1 L_e should satisfy $L_e \geq n_a + Mn_b$.) Let $\overline{L} = L_e$ in (19). The following steps are executed to implement a practical algorithm based on the earlier discussion in Secs. 2.1–2.4.

ALGORITHM I: Multistep Linear Predictors-Based Blind equalizer – MSLP Estimate all data-based correlations by sample averaging. Estimate noisefree correlations via (20). Use the steps in Sec. 2.4 to estimate $\mathbf{F}_i$ for $i = 0,1,\cdots,d$. Using these estimated values, implement the MMSE equalizer via (21) and (22).

ALGORITHM II: Linear Predictor-Based Blind equalizer – LP Here we will use (21) with $\mathbf{F}_i$ ($i = 0,1,\cdots,d$) estimated using the basic approach of [1] which utilizes one-step ahead linear predictors ($l = 1$). Although [1] derives all its results under the assumption of FIR column-reduced irreducible channels, its results extend to models that satisfy (**H1**)-(**H5**) by virtue of Theorem 1.

3. COMMON MINIMUM-PHASE ZEROS

Here the MIMO transfer function is

$$\mathcal{F}(z) = \mathcal{A}^{-1}(z)\mathcal{B}(z)\mathcal{B}_c(z), \quad \mathcal{B}_c(z) = \sum_{i=0}^{n_{bc}} \mathbf{B}_{ci}z^{-i} \quad (35)$$

where $\mathcal{B}(z)$ satisfies (**H2**) and $\mathcal{B}_c(z)$ is a finite-degree $M \times M$ polynomial that collects all the common zeros/factors of the subchannels. Assume that

(**H6**) Given model (35), $\det(\mathcal{B}_c(z)) \neq 0$ for $|z| \geq 1$.

Then while $\mathcal{A}^{-1}(z)\mathcal{B}(z)$ has a finite left-inverse [14, Thm. 2], $\mathcal{B}_c^{-1}(z)$ is IIR though causal under (**H6**). Therefore, we have (for $0 < \alpha < \infty$ and $0 < \beta < 1$)

$$\mathbf{w}(k) = \sum_{i=0}^{\infty} \mathbf{G}_i'\mathbf{s}(k-i) \text{ where } \|\mathbf{G}_i'\| \leq \alpha\beta^{|i|} \ \forall i. \quad (36)$$

Consider (6) with $\mathbf{e}(k|k-l)$ as in (7) but

$$\widehat{\mathbf{s}}(k|k-l) := \sum_{i=l}^{\infty} \mathbf{F}_i w(k-i). \quad (37)$$

Using (36) and (37) we have (for $0 < \alpha_1 < \infty$ and $0 < \beta_1 < 1$)

$$\widehat{\mathbf{s}}(k|k-l) = \sum_{n=l}^{\infty} \mathbf{C}_n\mathbf{s}(k-n) \text{ where } \|\mathbf{C}_n\| \leq \alpha_1\beta_1^{|n|} \ \forall n. \quad (38)$$

Mimicking Theorem 1, we have the following result.

Theorem 2. With $\mathcal{F}(z)$ obeying (35), under (**H1**)-(**H3**), (**H5**), (**H6**) and for $l = 1,2,\cdots$, $\{\mathbf{s}(k)\}$ can be decomposed as in (6), (7) and (37) such that (9) and (10) hold true. Furthermore, let $\widehat{\mathbf{s}}(k|k-l,k-Q) := (\mathbf{s}(k)|H_{k-l;Q,\cdots,Q}(\mathbf{s}))$ and $\mathbf{e}(k|k-l,k-Q) := \mathbf{s}(k) - \widehat{\mathbf{s}}(k|k-l,k-Q)$. Then

$$\lim_{Q\to\infty} E\{\|\mathbf{e}(k|k-l) - \mathbf{e}(k|k-l,k-Q)\|^2\} = 0 \bullet \quad (39)$$

Proof: The first part follows as in Theorem 1. It follows from [4, Chapter 1, Lemma 3.1.b] that

$$\lim_{Q\to\infty} E\{\|\widehat{\mathbf{s}}(k|k-l) - \widehat{\mathbf{s}}(k|k-l,k-Q)\|^2\} = 0. \quad (40)$$

Then (39) is immediate. □

Theorem 2 clearly implies that for Q 'large enough,' we can obtain $\mathbf{e}(k|k-l,k-Q)$ close enough to $\mathbf{e}(k|k-l)$. Therefore, the approach of Sec. 2 becomes applicable to the current case.

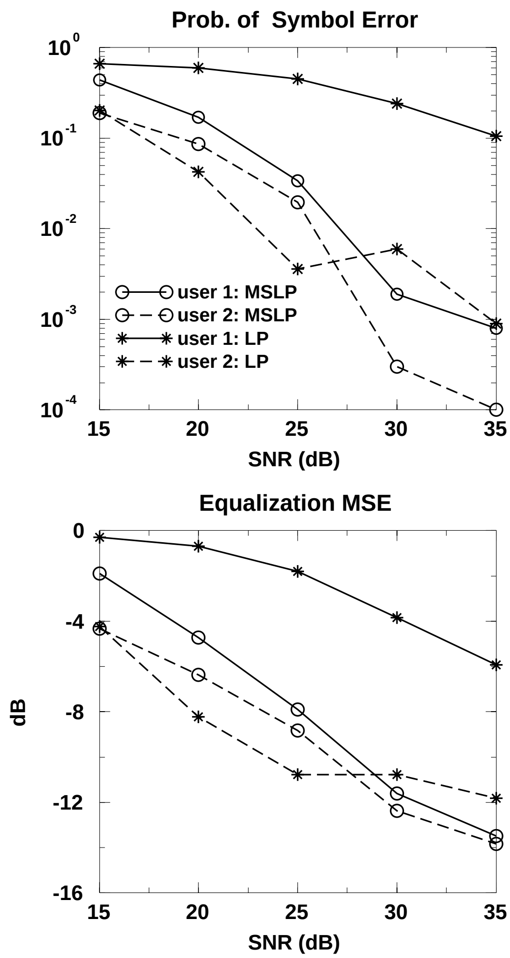

Fig. 1. Normalized MSE and probability of symbol detection error P_e for the two users after MMSE equalization with $d = 3$. Record length $T = 500$ symbols for equalizer design. Averages over 100 Monte Carlo runs. The designed equalizer was applied to record lengths of 3000 symbols for performance evaluation.

4. SIMULATION EXAMPLE

We consider a wireless communications scenario with two ($M = 2$) 4-QAM user signals arriving at a uniform linear array (half-wavelength spacing) of $N = 4$ sensors via a frequency selective multipath channel. The signaling pulse shape for both the users was a raised-cosine pulse with a roll-off factor of 0.2, the pulse being truncated to a length of $4T_s$ where T_s = symbol duration. The array measurements are assumed to be sampled at baud rate with sampling interval T_s seconds and the two sources have the same baud rate. The relative time delay τ (relative to the first arrival), the angle of arrival θ (in degrees w.r.t. the array broadside) and the relative attenuation factor (amplitude) α, (τ, θ, α), for the two sources were selected as:

$$w_1 : (0T_s, 40^\circ, 1.0),\ (0.3T_s, 20^\circ, 1.0),\ (0.6T_s, -20^\circ, 1.0)$$

$$w_2 : (0T_s, 10^\circ, 1.0),\ (1.1T_s, -15^\circ, 1.0),\ (1.6T_s, -1^\circ, 1.0). \quad (41)$$

Sampling of received signal at the array leads to a discrete-time MIMO FIR model $\mathcal{B}(z)$ with $N = 4$, $M = 2$ and $n_b = 4$ such that $\mathbf{B}_0 \neq 0$ and $\mathbf{B}_4 \neq 0$ (see Sec. 1). The effective MIMO channel was taken to be

$$\mathcal{F}(z) = \mathcal{B}(z)\mathcal{B}_c(z) \text{ where } \mathcal{B}_c(z) = (1 - 0.5z^{-1})I_2. \quad (42)$$

The part $\mathcal{B}_c(z)$ in (42) may represent some filtering at the transmitter or receiver, and it leads to a system with common zeros in the two subchannels.

An MMSE equalizer of length $L_e = 9$ (10 taps per subchannel, totaling 32 taps) was designed with a delay d =3 for each of the two sources. In order to apply algorithms MSLP and LP, we picked $\overline{L} = L_e = 9$. Fig. 1 shows the results of simulations for a record length of T =500 symbols. It is seen that the MSLP approach outperforms the LP approach for user 1 whereas their performances are comparable for user 2. Also presence of a common (minimum-phase) zero has not caused any problems.

5. REFERENCES

[1] A. Gorokhov, P. Loubaton and E. Moulines, "Second order blind equalization in multiple input multiple output FIR systems: A weighted least squares approach," in *Proc. 1996 ICASSP*, pp. 2415-2418, Atlanta, GA, May 7-10, 1996.

[2] K. Abed-Meraim, E. Moulines and P. Loubaton, "Prediction error method for second-order blind identification," *IEEE Trans. Signal Processing*, vol. SP-45, pp. 694-705, March 1997.

[3] Special Issue, *IEEE Trans. Signal Processing*, vol. SP-45, Jan. 1997.

[4] P.E. Caines, *Linear Stochastic Systems.* New York: Wiley, 1988.

[5] T. Kailath, *Linear Systems.* Englewood Cliffs, NJ: Prentice-Hall, 1980.

[6] N. Delfosse and P. Loubaton, "Adaptive blind separation of convolutive mixtures," in *Proc. 1996 ICASSP*, pp. 2940-2943, Atlanta, GA, May 7-10, 1996.

[7] J.K. Tugnait, "Blind spatio-temporal equalization and impulse response estimation for MIMO channels using a Godard cost function," *IEEE Trans. Signal Processing*, vol. SP-45, pp. 268-271, Jan. 1997.

[8] J.F. Cardoso and A. Souloumiac, "Blind beamforming for non-Gaussian signals," *IEE Proc.-F, Radar & Signal Processing*, vol. 140, pp. 362-370, Dec. 1993.

[9] D. Slock, "Blind fractionally-spaced equalization, perfect reconstruction filter banks and multichannel linear prediction," in *Proc. 1994 IEEE ICASSP*, pp. IV:585-588, Adelaide, Australia, May 1994.

[10] L. Tong, G. Xu and T. Kailath, "A new approach to blind identification and equalization of multipath channels," *IEEE Trans. Information Theory*, vol. IT-40, pp. 340-349, March 1994.

[11] E. Moulines, P. Duhamel, J. Cardoso and S. Mayrargue, "Subspace methods for blind identification of multichannel FIR filters," *IEEE Trans. Signal Proc.*, vol. SP-43, pp. 516-525, Feb. 1995.

[12] Z. Ding, "Matrix outer-product decomposition method for blind multiple channel identification," *IEEE Trans. Signal Proc.*, vol. SP-45, pp. 3053-3061, Dec. 1997.

[13] D. Gesbert and P. Duhamel, "Robust blind channel identification and equalization based on multi-step predictors," in *Proc. 1997 ICASSP*, pp. 3621-3624, April 21-24, 1997.

[14] J.K. Tugnait, "On multi-step linear predictors for M.I.M.O. F.I.R./I.I.R. channels and related blind equalization," in *Proc. 1998 IEEE DSP Workshop*, Paper #110, Bryce Canyon, Utah, Aug. 1998.