BLIND CHANNEL ESTIMATION AND 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

Channel estimation and blind equalization of MIMO (multiple-input multiple-output) communications channels is considered using primarily the second-order statistics of the data. We consider estimation of (partial) channel impulse response and design of finite-length MMSE (minimum mean-square error) blind equalizers. The basis of the approach is the design of a zero-forcing equalizer that whitens the noise-free data. We allow infinite impulse response (IIR) channels. Moreover, the multichannel transfer function need not be column-reduced. Our approaches also work when the "subchannel" transfer functions have common zeros so long as the common zeros are minimum-phase zeros. The channel length or model orders need not be known. The sources are recovered up to a unitary mixing matrix and are further 'unmixed' using higher-order statistics of the data. An illustrative simulation example is provided.

1. INTRODUCTION

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

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

where $\mathbf{y}(k) = [y_1(k) \vdots y_2(k) \vdots \cdots \vdots y_N(k)]^T$, similarly for $\mathbf{s}(k)$, $\mathbf{w}(k)$ and $\mathbf{n}(k)$, and 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

$$\mathcal{F}(z) \;:=\; \sum_{l=0}^{\infty} \mathbf{F}_l z^{-l} \;=\; \mathcal{A}^{-1}(z)\mathcal{B}(z), \qquad (1-2)$$

$$\mathcal{A}(z) = I + \sum_{i=1}^{n_a} \mathbf{A}_i z^{-i} \;\text{ and }\; \mathcal{B}(z) \;=\; \sum_{i=0}^{n_b} \mathbf{B}_i z^{-i}. \quad (1-3)$$

We allow all of the above variables to be complex-valued.

Such models arise in several useful baseband-equivalent digital communications and other applications [1]-[6],[8], [10]-[12],[14]. In these applications one of the objectives is to recover the inputs $\mathbf{w}(k)$ given the noisy measurements but not given the knowledge of the system transfer function. A large number of papers (see [4],[5],[10],[11],[14]) have concentrated on a two-step procedure: first estimate the channel impulse response (IR) and then design an equalizer using the estimated channel. A fundamental restriction in these works is that the channel is FIR with no common zeros among the various subchannels. A few (see [1]and [12], e.g.) have proposed direct design of the equalizer bypassing channel estimation. Still they assume FIR (SIMO) channels with no common zeros.

In this paper we allow IIR channels. The MIMO channel does not have to be column-reduced. We will also allow common zeros so long as they are minimum-phase. Finally, in the presence of nonminimum-phase common zeros, our proposed approach equalizes the spectrally-equivalent minimum-phase counterpart of $\mathcal{F}(z)$; it does not "fall apart" unlike quite a few existing approaches. Our proposed approach extends the SIMO results of [1] and [8] to MIMO channels.

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

2. PRELIMINARIES

2.1. FIR Inverses

Assume the following:

- **(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 [7, Sec. 6.3].
- **(H3)** $\det(\mathcal{A}(z)) \neq 0$ for $|z| \geq 1$.

It has been shown in [6] (using some results from [2]) that under **(H1)-(H3)** there exists a finite degree left-inverse (not necessarily unique) of $\mathcal{F}(z)$:

$$\mathcal{G}(z)\mathcal{F}(z) \;=\; I_M \qquad (2-1)$$

where $\mathcal{G}(z)$ is $M \times N$ given by

$$\mathcal{G}(z) \;=\; \sum_{l=0}^{L_e} \mathbf{G}_l z^{-l} \;\text{ for any }\; L_e \geq n_a + (2M-1)n_b - 1. \qquad (2-2)$$

2.2. Linear Innovations Representation

Assume further the following:

- **(H4)** 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 $E\{\mathbf{w}(k)\mathbf{w}^{\mathcal{H}}(k)\} = I_M$ where I_M is the $M \times M$ identity matrix and the superscript $\mathcal{H}$ is the Hermitian operation.

Lemma 1. Under **(H1)**-**(H4)**, $\{\mathbf{s}(k)\}$ may be represented as

$$\mathbf{s}(k) \;=\; -\sum_{i=1}^{K} \mathbf{D}_i \mathbf{s}(k-i) \;+\; I_s(k) \qquad (2-3)$$

where $K \leq n_a + Mn_b$, $\mathbf{D}_i$'s are some $N \times N$ matrices and $\{I_s(k)\}$ is a zero-mean white $N \times 1$ random sequence (linear innovations for $\{\mathbf{s}(k)\}$) with

$$I_s(k) \;=\; \mathbf{F}_0 \mathbf{w}(k). \;\bullet \qquad (2-4)$$

Proof: See [16]. □

It follows from (1-1) and Lemma 1 that

$$\mathcal{D}(z)\mathbf{s}(k) = \mathcal{D}(z)\mathcal{F}(z)\mathbf{w}(k) = \mathbf{F}_0\mathbf{w}(k) \qquad (2-5)$$

where $\mathcal{D}(z) = I_N + \sum_{i=1}^{K} \mathbf{D}_i z^{-i}$. Since $\mathbf{w}(k)$ is full-rank and white, it follows from (2-5) that

$$\mathcal{D}(z)\mathcal{F}(z) = \mathbf{F}_0 \quad \Rightarrow \quad \left(\mathbf{F}_0^{\mathcal{H}}\mathbf{F}_0\right)^{-1}\mathbf{F}_0^{\mathcal{H}}\mathcal{D}(z)\mathcal{F}(z) = I_M. \qquad (2-6)$$

Clearly the $M \times N$ polynomial matrix $\mathcal{G}(z) := \left(\mathbf{F}_0^{\mathcal{H}}\mathbf{F}_0\right)^{-1}\mathbf{F}_0^{\mathcal{H}}\mathcal{D}(z)$ is of degree $K \le n_a + Mn_b$ and it is a left inverse to $\mathcal{F}(z)$. This result is summarized below.

Lemma 2. Under (**H1**)-(**H4**), there exists an integer $K \le n_a + Mn_b$ and a polynomial matrix $\mathcal{G}(z) = \sum_{i=0}^{K}\mathbf{G}_i z^{-i}$ of degree K such that $\mathcal{G}(z)\mathcal{F}(z) = I_M$. •

Lemma 3. Let $\mathcal{R}_{ssL_e}$ denote a $[N(L_e+1)]\times[N(L_e+1)]$ matrix with its ij-th block element as $\mathbf{R}_{ss}(j-i) = E\{\mathbf{s}(k+j-i)\mathbf{s}^{\mathcal{H}}(k)\}$. Then under (**H1**)-(**H4**), $\rho(\mathcal{R}_{ssL_e}) \le NL_e + M$ for $L_e \ge n_a + Mn_b$ where $\rho(A)$ denotes the rank of A. •
Sketch of proof: It follows from Lemma 1 and (2-3) that

$$\begin{bmatrix} I_N & \mathbf{D}_1 & \cdots & \mathbf{D}_{n_a+Mn_b} & 0 & \cdots & 0 \end{bmatrix}\mathcal{R}_{ssL_e} = \begin{bmatrix} \mathbf{F}_0\mathbf{F}_0^{\mathcal{H}} & 0 & \cdots & 0 \end{bmatrix}. \qquad (2-7)$$

Apply Sylvester's inequality [7, p. 655] to (2-7) to deduce the desired result. □

3. EQUALIZATION: NO COMMON ZEROS

Assume that (**H1**)-(**H4**) hold true. In addition assume the following regarding the measurement noise:

(**H5**) $\{\mathbf{n}(k)\}$ is zero-mean Gaussian with $E\{\mathbf{n}(k+\tau)\mathbf{n}^{\mathcal{H}}(k)\} = \sigma_n^2 I_N\delta(\tau)$.

3.1. Zero-Delay Zero-Forcing Blind Equalizer

Using (1-2), (2-1) and (2-2), we have

$$\sum_{l=0}^{\infty}\mathbf{G}_{m-l}\mathbf{F}_l = \begin{cases} I_M, & m=0 \\ 0, & m=1,2,\cdots, \end{cases} \qquad (3-1)$$

leading to

$$\begin{bmatrix} \mathbf{G}_0 & \mathbf{G}_1 & \cdots & \mathbf{G}_{L_e} \end{bmatrix}\overline{\mathcal{S}} = \begin{bmatrix} 1 & 0 & \cdots & \cdots \end{bmatrix} \qquad (3-2)$$

where $\overline{\mathcal{S}}$ is the $(N(L_e+1))\times\infty$ matrix given by

$$\overline{\mathcal{S}} = \begin{bmatrix} \mathbf{F}_0 & \mathbf{F}_1 & \mathbf{F}_2 & \mathbf{F}_3 & \cdots & \cdots & \cdots \\ 0 & \mathbf{F}_0 & \mathbf{F}_1 & \mathbf{F}_2 & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & & & \\ 0 & 0 & \cdots & 0 & \mathbf{F}_0 & \mathbf{F}_1 & \cdots \end{bmatrix}. \qquad (3-3)$$

Let $\overline{\mathcal{S}}^{\#}$ denote the pseudoinverse of $\overline{\mathcal{S}}$. By [15, Prop. 1], $\overline{\mathcal{S}}^{\#} = \overline{\mathcal{S}}^{\mathcal{H}}(\overline{\mathcal{S}}\,\overline{\mathcal{S}}^{\mathcal{H}})^{\#}$. Then the minimum norm solution to the FIR equalizer is given by [15, Sec. 6.11]

$$\begin{bmatrix} \mathbf{G}_0 & \mathbf{G}_1 & \cdots & \mathbf{G}_{L_e} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_0^{\mathcal{H}} & 0 & \cdots & 0 \end{bmatrix}(\overline{\mathcal{S}}\,\overline{\mathcal{S}}^{\mathcal{H}})^{\#}. \qquad (3-4)$$

In a fashion similar to $\mathcal{R}_{ssL_e}$ in Lemma 2, let $\mathcal{R}_{yyL_e}$ denote a $[N(L_e+1)]\times[N(L_e+1)]$ matrix with its ij-th block element as $\mathbf{R}_{yy}(j-i) = E\{\mathbf{y}(k+j-i)\mathbf{y}^{\mathcal{H}}(k)\}$; define similarly $\mathcal{R}_{nnL_e}$ pertaining to the additive noise. Carry out an eigendecomposition of $\mathcal{R}_{yyL_e}$. Then the smallest $N-M$ eigenvalues of $\mathcal{R}_{yyL_e}$ equal σ_n^2 because under (**H1**)-(**H4**), $\rho(\mathcal{R}_{ssL_e}) \le NL_e + M$ whereas $\rho(\mathcal{R}_{nnL_e}) = NL_e + N = \rho(\mathcal{R}_{yyL_e})$. Thus 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 $\widehat{\mathcal{R}}_{yyL_e}$, the data-based consistent estimate of $\mathcal{R}_{yyL_e}$.

Under (**H4**) and (**H5**),

$$(\overline{\mathcal{S}}\,\overline{\mathcal{S}}^{\mathcal{H}}) = \mathcal{R}_{ssL_e} = \mathcal{R}_{yyL_e} - \mathcal{R}_{nnL_e} = \mathcal{R}_{yyL_e} - \sigma_n^2 I. \qquad (3-5)$$

Thus, $(\overline{\mathcal{S}}\,\overline{\mathcal{S}}^{\mathcal{H}})$ can be estimated from noisy data. However, we don't know $\mathbf{F}_0$. To this end, we seek an $N\times N$ FIR filter $\mathcal{G}_a(z) := \sum_{i=0}^{L_e}\mathbf{G}_{ai}z^{-i}$ satisfying

$$\begin{bmatrix} \mathbf{G}_{a0} & \mathbf{G}_{a1} & \cdots & \mathbf{G}_{aL_e} \end{bmatrix} = \begin{bmatrix} I_N & 0 & \cdots & 0 \end{bmatrix}\mathcal{R}_{ssL_e}^{\#}. \qquad (3-6)$$

Comparing (3-4) and (3-6) it follows that

$$\begin{bmatrix} \mathbf{G}_0 & \mathbf{G}_1 & \cdots & \mathbf{G}_{L_e} \end{bmatrix} = \mathbf{F}_0^{\mathcal{H}}\begin{bmatrix} \mathbf{G}_{a0} & \mathbf{G}_{a1} & \cdots & \mathbf{G}_{aL_e} \end{bmatrix} \qquad (3-7)$$

leading to

$$\sum_{i=0}^{L_e}\mathbf{G}_i z^{-i} =: \mathcal{G}(z) = \mathbf{F}_0^{\mathcal{H}}\mathcal{G}_a(z). \qquad (3-8)$$

In practice, therefore, we apply $\mathcal{G}_a(z)$ to the data leading to

$$\mathbf{v}(k) := \mathcal{G}_a(z)\mathbf{y}(k) = \mathbf{v}_s(k) + \mathcal{G}_a(z)\mathbf{n}(k) \qquad (3-9)$$

such that

$$\mathbf{F}_0^{\mathcal{H}}\mathbf{v}_s(k) = \mathbf{w}(k) \qquad (3-10)$$

where

$$\mathbf{v}_s(k) := \mathcal{G}_a(z)[\mathbf{y}(k) - \mathbf{n}(k)] = \mathcal{G}_a(z)\mathbf{s}(k). \qquad (3-11)$$

In (3-10) $\{\mathbf{w}(k)\}$ is a white $M-$vector sequence (by assumption (**H4**)), however, $\{\mathbf{v}_s(k)\}$ is not necessarily a white vector sequence. Given the second-order statistics of $\{\mathbf{v}_s(k)\}$, how does one estimate $\mathbf{F}_0$ so that $\{\mathbf{w}(k)\}$ satisfying (**H4**) is recovered? We need to have $\mathbf{R}_{ww}(\tau) := E\{\mathbf{w}(k+\tau)\mathbf{w}^{\mathcal{H}}(k)\} = 0$ for $|\tau| \ne 0$ and $\rho(\mathbf{R}_{ww}(0)) = M$. By (3-10), $\mathbf{R}_{ww}(\tau) = \mathbf{F}_0^{\mathcal{H}}\mathbf{R}_{v_sv_s}(\tau)\mathbf{F}_0$ where $\mathbf{R}_{v_sv_s}(\tau) := E\{\mathbf{v}_s(k+\tau)\mathbf{v}_s^{\mathcal{H}}(k)\}$. Define ($L>0$ is some large integer)

$$\overline{R}_{v_sv_s} := [\mathbf{R}_{v_sv_s}^T(-1)\ \ \mathbf{R}_{v_sv_s}^T(-2)\ \ \cdots\ \ \mathbf{R}_{v_sv_s}^T(-L) \vdots Q^*]^T \qquad (3-12)$$

where $Q = [\mathbf{q}_{r+1}\vdots\cdots\vdots\mathbf{q}_N]$, $r = \rho(\mathbf{R}_{v_sv_s}(0))$ with $M \le r \le N$, $\mathbf{q}_i$'s $(r+1 \le i \le N)$ are orthonormal eigenvectors of $\mathbf{R}_{v_sv_s}(0)$ corresponding to zero eigenvalues, and the symbol $*$ denotes the complex conjugation operation. Note that if $r = N$, Q is omitted from (3-12). Let $\mathbf{F}_{0s}$ and $\mathbf{F}_{0n}$ be the orthogonal projections of $\mathbf{F}_0$ onto the $r-$dimensional 'signal subspace' (range space) and $(N-r)-$dimensional 'noise subspace' (null space), respectively, of $\mathbf{R}_{v_sv_s}(0)$. By (3-10), $\rho(\mathbf{F}_{0s}) = M$. Then $\mathbf{F}_0 = \mathbf{F}_{0s} + \mathbf{F}_{0n}$ with $\mathbf{F}_{0s}^{\mathcal{H}}\mathbf{F}_{0n} = 0$ and $\mathbf{R}_{v_sv_s}(0)\mathbf{F}_{0n} = 0$. It then follows that $E\{\mathbf{F}_{0n}^{\mathcal{H}}\mathbf{v}_s(k)\mathbf{v}_s^{\mathcal{H}}(k)\mathbf{F}_{0n}\} = \mathbf{F}_{0n}^{\mathcal{H}}\mathbf{R}_{v_sv_s}(0)\mathbf{F}_{0n} = 0$; hence, $\mathbf{F}_{0n}^{\mathcal{H}}\mathbf{v}_s(k) = 0$ with probability one (w.p.1) and (cf. (3-10))

$$\mathbf{F}_{0s}^{\mathcal{H}}\mathbf{v}_s(k) = \mathbf{w}(k). \qquad (3-13)$$

Lemma 4. $\overline{R}_{v_sv_s}$ is rank deficient for any $L \ge 1$ such that $\overline{R}_{v_sv_s}\mathbf{F}_{0s} = 0$ and $\rho(\mathbf{F}_{0s}) = M$. •
Proof: By construction $Q^{\mathcal{H}}\mathbf{F}_{0s} = 0$ as columns of Q span the noise-subspace of $\mathbf{R}_{v_sv_s}(0)$ and $\mathbf{F}_{0s}$ is the orthogonal projection of $\mathbf{F}_0$ onto the $r-$dimensional signal subspace of $\mathbf{R}_{v_sv_s}(0)$. Furthermore, we have

$$\mathbf{R}_{wv_s}(\tau) = E\{\mathbf{w}(k+\tau)\mathbf{v}_s^{\mathcal{H}}(k)\} = 0 \quad \forall\tau \ge 1 \qquad (3-14)$$

because $\mathbf{v}_s(k)$ is obtained by causal filtering of $\mathbf{y}(k)$, hence of $\mathbf{w}(k)$. Using (3-13) in (3-14) it then follows that there exists a $N\times M$ $\mathbf{F}_{0s} \ne 0$ such that

$$\mathbf{F}_{0s}^{\mathcal{H}}\mathbf{R}_{v_sv_s}(\tau) = 0\ \forall\tau \ge 1 \ \Rightarrow\ \mathbf{R}_{v_sv_s}(-\tau)\mathbf{F}_{0s} = 0\ \forall\tau \ge 1. \qquad (3-15)$$

The desired result is then immediate. □

Pick a $N \times M$ column-vector $\mathbf{H}_0$ to equal the rightmost M right singular vectors in a singular-value decomposition (SVD) $\overline{R}_{v_s v_s} = U \Sigma V^{\mathcal{H}}$, i.e. the right singular vectors corresponding to the M smallest singular values. Therefore, $\rho(\mathbf{H}_0) = M$. Then since ideally the M smallest singular values of $\overline{R}_{v_s v_s}$ are zero, we have $\mathbf{H}_0^{\mathcal{H}} \mathbf{R}_{v_s v_s}(-\tau)\mathbf{H}_0 = 0$ for $\tau = 1, 2, \cdots, L$. This, in turn, implies that

$$\left(\mathbf{H}_0^{\mathcal{H}} \mathbf{R}_{v_s v_s}(-\tau)\mathbf{H}_0\right)^{\mathcal{H}} \;=\; 0 \;\; \text{for} \;\;\; \tau = 1, 2, \cdots, L. \quad (3-16)$$

Moreover $Q^{\mathcal{H}}\mathbf{H}_0 = 0$. An eigendecomposition yields

$$\mathbf{R}_{v_s v_s}(0) = \sum_{i=1}^{r} \sigma_i \mathbf{q}_i \mathbf{q}_i^{\mathcal{H}} = \overline{Q}\Sigma\overline{Q}^{\mathcal{H}} \qquad (3-17)$$

where $\overline{Q} = [\mathbf{q}_1 \vdots \cdots \vdots \mathbf{q}_r]$, $\mathbf{q}_i$'s $(1 \le i \le r)$ are orthonormal eigenvectors of $\mathbf{R}_{v_s v_s}(0)$ corresponding to non-zero eigenvalues σ_i's and $\Sigma = \text{diag}\{\sigma_1, \cdots, \sigma_r\}$. Thus columns of $\overline{Q}$ span the signal subspace of $\mathbf{R}_{v_s v_s}(0)$ and the columns of Q span the noise subspace of $\mathbf{R}_{v_s v_s}(0)$. Since $Q^{\mathcal{H}}\mathbf{H}_0 = 0$, we have $\mathbf{H}_0 = \overline{Q}\mathbf{C}$ where $\mathbf{C}$ is $r \times M$ and $\rho(\mathbf{C}) = M$. Then $\mathbf{H}_0^{\mathcal{H}}\mathbf{R}_{v_s v_s}(0)\mathbf{H}_0 = \widetilde{\mathbf{C}}^{\mathcal{H}}\widetilde{\mathbf{C}}$ where $\widetilde{\mathbf{C}} = \Sigma^{1/2}\mathbf{C}$ and $\rho(\widetilde{\mathbf{C}}) = M$. Finally, by result R11 on p. 261 of [9], $\rho(\widetilde{\mathbf{C}}^{\mathcal{H}}\widetilde{\mathbf{C}}) = \rho(\widetilde{\mathbf{C}}\widetilde{\mathbf{C}}^{\mathcal{H}})$, and therefore, by Sylvester's inequality [7, p. 655], $\rho(\mathbf{H}_0^{\mathcal{H}}\mathbf{R}_{v_s v_s}(0)\mathbf{H}_0) = M$.

The overall system with $\mathbf{w}(k)$ as input and $\mathbf{H}_0^{\mathcal{H}}\mathbf{v}_s(k)$ as output has the transfer function

$$\mathbf{H}_0^{\mathcal{H}}\mathcal{G}_a(z)\mathcal{A}^{-1}(z)\mathcal{B}(z)$$

$$= [\det(\mathcal{A}(z))I_M]^{-1}\mathbf{H}_0^{\mathcal{H}}\mathcal{G}_a(z)\text{adj}(\mathcal{A}(z))\mathcal{B}(z) \qquad (3-18)$$

and therefore, is an autoregressive moving average (ARMA) model with autoregressive (AR) order no more than Nn_a and moving average (MA) order no more than $n_b + L_e + (N-1)n_a$, denoted by ARMA$(Nn_a, n_b + L_e + (N-1)n_a)$. Therefore, it follows from (3-16) that $\mathbf{H}_0^{\mathcal{H}}\mathbf{v}_s(k)$ is zero-mean white if $L \ge n_b + L_e + (N-1)n_a$. Moreover, since $\mathbf{v}_s(k)$ is obtained by causal filtering of $\mathbf{w}(k)$, it follows that $\mathbf{H}_0^{\mathcal{H}}\mathbf{v}_s(k) = \sum_{i=0}^{\infty}\mathbf{P}_i\mathbf{w}(k-i)$ where $M \times M$ $\mathcal{P}(z) = \sum_{i=0}^{\infty}\mathbf{P}_i z^{-i}$ is stable satisfying $\mathcal{P}(e^{j\omega})\mathcal{P}^{\mathcal{H}}(e^{j\omega}) = I_M \; \forall\omega$ (allpass). Therefore, we have $E\{\mathbf{H}_0^{\mathcal{H}}\mathbf{v}_s(k+\tau)\mathbf{w}^{\mathcal{H}}(k)\} = \mathbf{P}_\tau$ for $\tau \ge 0$, and using (3-13), we have $E\{\mathbf{H}_0^{\mathcal{H}}\mathbf{v}_s(k+\tau)\mathbf{w}^{\mathcal{H}}(k)\} = \mathbf{H}_0^{\mathcal{H}}\mathbf{R}_{v_s v_s}(\tau)\mathbf{F}_{0s} = 0$ (by construction of $\mathbf{H}_0$) for $\tau \ge 1$. Thus $\mathbf{P}_\tau = 0$ for $\tau \ge 1$. Therefore, we have

$$\mathbf{H}_0^{\mathcal{H}}\mathbf{v}_s(k) \;=\; \mathbf{P}_0\mathbf{w}(k) \;\;\; \text{such that} \;\;\; \rho(\mathbf{P}_0) = M. \quad (3-19)$$

By (3-11) and (3-19), we have

$$\mathbf{H}_0^{\mathcal{H}}\mathcal{G}_a(z)\mathbf{y}(k) = \mathbf{P}_0\mathbf{w}(k) + \mathbf{H}_0^{\mathcal{H}}\mathcal{G}_a(z)\mathbf{n}(k) \;\ni\; \rho(\mathbf{P}_0) = M. \quad (3-20)$$

Since $\mathbf{P}_0\mathbf{U}\mathbf{U}^{\mathcal{H}}\mathbf{P}_0^{\mathcal{H}} = \mathbf{P}_0\mathbf{P}_0^{\mathcal{H}}$ for any unitary matrix $\mathbf{U}$, one can not uniquely determine $\mathbf{P}_0$ from (3-20) given second-order statistics of data $\mathbf{y}(k)$, and knowledge (estimates) of $\mathbf{H}_0$, filter $\mathcal{G}_a(z)$ and noise variance σ_n^2 [13]. One has to exploit higher-order statistics (HOS) of data. The model (3-20) is an instantaneous mixture model [13], therefore, any existing method may be applied to estimate $\mathbf{P}_0$ given (3-20). In this paper we have used the joint diagonalization procedure of [13]. The required estimate of $\mathbf{P}_0$ is obtained as

$$\mathbf{P}_0 = \mathbf{W}^{-1}\mathbf{U} \qquad (3-21)$$

where $\mathbf{W}$ "diagonalizes" $\mathbf{H}_0^{\mathcal{H}}\mathbf{R}_{v_s v_s}(0)\mathbf{H}_0$ into an identity matrix and $\mathbf{U}$ is a unitary matrix obtained via the joint diagonalization procedure of [13] using fourth-order cumulants of $\mathbf{W}\mathbf{H}_0^{\mathcal{H}}\mathcal{G}_a(z)\mathbf{y}(k)$. Let $\mathbf{l}_i$ $(i = 1, 2, \cdots, M)$ denote the orthonormal eigenvectors of $\mathbf{H}_0^{\mathcal{H}}\mathbf{R}_{v_s v_s}(0)\mathbf{H}_0$ with the corresponding eigenvalues γ_i's. Set $\mathcal{L} = [\mathbf{l}_1 \vdots \cdots \vdots \mathbf{l}_M]$ and $, = diag(\gamma_1, \cdots, \gamma_M)$. Then

$$\mathbf{W} \;=\; ,^{-1/2}\mathcal{L}^{\mathcal{H}} \qquad (3-22)$$

diagonalizes $\mathbf{H}_0^{\mathcal{H}}\mathbf{R}_{v_s v_s}(0)\mathbf{H}_0$ to an identity matrix: $\mathbf{W}\left(\mathbf{H}_0^{\mathcal{H}}\mathbf{R}_{v_s v_s}(0)\mathbf{H}_0\right)\mathbf{W}^{\mathcal{H}} = I_M$. Finally, we have

$$\overline{\mathbf{H}}_0^{\mathcal{H}}\mathbf{v}_s(k) \;=:\; \mathbf{w}(k) \;\;\; \text{where} \;\;\; \overline{\mathbf{H}}_0 = \mathbf{H}_0\mathbf{W}^{\mathcal{H}}\mathbf{U}. \quad (3-23)$$

Remark 1. Using (3-11) and (3-23) it follows that $\overline{\mathbf{H}}_0^{\mathcal{H}}\left[\sum_{i=0}^{L_e}\mathbf{G}_{ai}\mathbf{s}(k-i)\right] = \mathbf{w}(k)$. The filter $\overline{\mathbf{H}}_0^{\mathcal{H}}\mathcal{G}_a(z)$ whitens the noise-free received signal. Moreover, the derivation of this filter was based upon whitening of $\overline{\mathbf{H}}_0^{\mathcal{H}}\mathcal{G}_a(z)\mathbf{s}(k)$. These considerations motivate the name a whitening approach for the proposed technique. Our approach is far more structured and different than that of [12]. □

3.2. MMSE Equalizer with Delay d

Using the orthogonality principle, the MMSE equalizer of length $L_e + 1$ to estimate $\mathbf{w}(k-d)$ $(d \ge 0)$ based upon $\mathbf{y}(n)$, $n = k, k-1, \cdots, k-L_e$, satisfies

$$\begin{bmatrix} \overline{\mathbf{G}}_{d,0} & \overline{\mathbf{G}}_{d,1} & \cdots & \overline{\mathbf{G}}_{d,L_e} \end{bmatrix} =$$

$$\begin{bmatrix} \mathbf{F}_d^{\mathcal{H}} & \mathbf{F}_{d-1}^{\mathcal{H}} & \cdots & \mathbf{F}_0^{\mathcal{H}} & 0 & \cdots & 0 \end{bmatrix}\mathcal{R}_{yyL_e}^{-1}. \quad (3-24)$$

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. Here the discussion of Sec. 3.1 becomes relevant. From (3-11) and (3-23) we have

$$\overline{\mathbf{H}}_0^{\mathcal{H}}\mathbf{v}_s(n) = \sum_{i=0}^{L_e}\overline{\mathbf{H}}_0^{\mathcal{H}}\mathbf{G}_{ai}\mathbf{s}(n-i). \qquad (3-25)$$

Using (3-25) and taking expectations we have

$$E\{\mathbf{s}(n)\mathbf{v}_s^{\mathcal{H}}(n-\tau)\}\overline{\mathbf{H}}_0 = \sum_{i=0}^{L_e}\mathbf{R}_{ss}(\tau+i)\mathbf{G}_{ai}^{\mathcal{H}}\overline{\mathbf{H}}_0. \quad (3-26)$$

Using (1-1), (1-2) and (3-23) we have

$$E\{\mathbf{s}(n)\mathbf{v}_s^{\mathcal{H}}(n-\tau)\}\overline{\mathbf{H}}_0 = \mathbf{F}_\tau. \qquad (3-27)$$

Hence, we have from (3-26) and (3-27)

$$\mathbf{F}_\tau^{\mathcal{H}} = \overline{\mathbf{H}}_0^{\mathcal{H}}\sum_{i=0}^{L_e}\mathbf{G}_{ai}\mathbf{R}_{ss}^{\mathcal{H}}(\tau+i). \qquad (3-28)$$

Substitute the results of (3-28) for $\tau = 0, 1, \cdots, d$ in (3-24) to complete the design. The MMSE estimate $\widehat{\mathbf{w}}(t-d)$ of $\mathbf{w}(t-d)$ is then given by $\widehat{\mathbf{w}}(t-d) = \sum_{i=0}^{L_e}\overline{\mathbf{G}}_{d,i}\mathbf{y}(t-i)$.

4. COMMON MINIMUM-PHASE ZEROS

Here the MIMO transfer function is

$$\mathcal{F}(z) \;=\; \mathcal{A}^{-1}(z)\mathcal{B}(z)\mathcal{B}_c(z), \quad \mathcal{B}_c(z) \;=\; \sum_{i=0}^{n_{bc}}\mathbf{B}_{ci}z^{-i} \qquad (4-1)$$

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 (4-1), $\det(\mathcal{B}_c(z)) \neq 0$ for $|z| \geq 1$.

Then while $\mathcal{A}^{-1}(z)\mathcal{B}(z)$ has a finite left-inverse, $\mathcal{B}_c^{-1}(z)$ is IIR though causal under (**H6**). Then (3-2) holds true approximately for "large" L_e, the approximation getting better with increasing L_e. Similarly Lemmas 1 and 2 hold true approximately for "large" K and Lemma 3 also holds true approximately for $L_e \geq K$. Note also that $\rho(\mathbf{F}_0) = M$ where $\mathbf{F}_0 = \mathbf{B}_0\mathbf{B}_{c0}$ since, by (**H6**), $\rho(\mathbf{B}_{c0}) = M$ (evaluate $\det(\mathcal{B}_c(z))$ at $z = \infty$). It is then readily seen that the developments of Sec. 3 apply to the current case also.

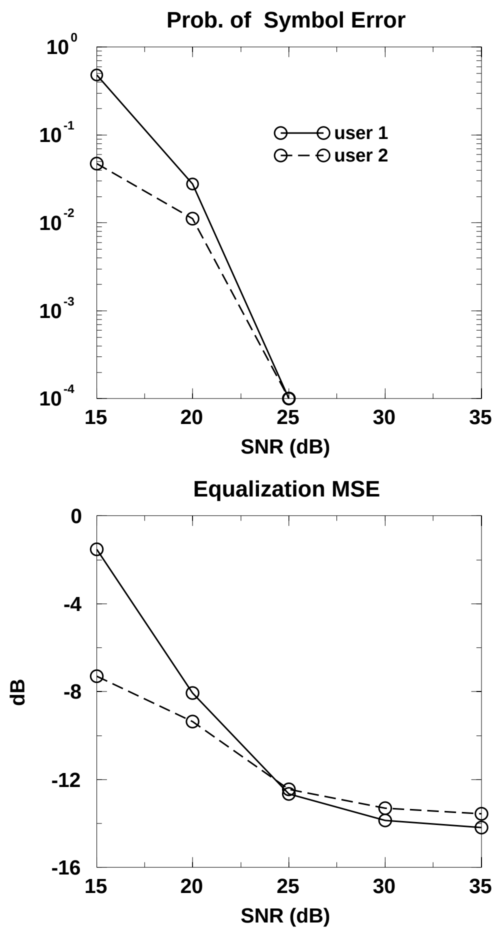

Fig. 1. Normalized MSE and probability of symbol detection error P_e for the two users after MMSE equalization with $d = 3$. Record length $T = 500$ symbols for equalizer design. Averages over 100 Monte Carlo runs. The designed equalizer was applied to record lengths of 3000 symbols for performance evaluation.

5. 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:

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

$$w_2 : (0T_s, 10^\circ, 1.0),\ (1.1T_s, -15^\circ, 1.0),\ (1.6T_s, -1^\circ, 1.0). \qquad (5-1)$$

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

$$\mathcal{F}(z) = \mathcal{B}(z)\mathcal{B}_c(z) \text{ where } \mathcal{B}_c(z) = (1 - 0.5z^{-1})I_2. \qquad (5-2)$$

The part $\mathcal{B}_c(z)$ in (5-2) 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 $L_e = 7$ (8 taps per subchannel, totaling 32 taps) was designed with a delay d =3 for each of the two sources. Fig. 1 shows the results of simulations for a record length of T =500 symbols. It is seen that the proposed approach works quite well. Also presence of a common (minimum-phase) zero has not caused any problems.

6. REFERENCES

[1] G.B. Giannakis and S.D. Halford, *IEEE Trans. Signal Proc.*, vol. SP-45, pp. 2277-2292, Sept. 1997.

[2] J.K. Tugnait, *IEEE Trans. Signal Processing*, vol. SP-45, pp. 268-271, Jan. 1997.

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

[4] L. Tong, G. Xu and T. Kailath, *IEEE Trans. Information Theory*, vol. IT-40, pp. 340-349, March 1994.

[5] E. Moulines, P. Duhamel, J. Cardoso and S. Mayrargue, *IEEE Trans. Signal Proc.*, vol. SP-43, pp. 516-525, Feb. 1995.

[6] J.K. Tugnait, in *Proc. 30th Annual Asilomar Conf. Signals Systems Computers*, pp. 295-299, Pacific Grove, CA, Nov. 1996.

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

[8] B. Huang & J.K. Tugnait, in *Proc. 1998 IEEE ICASSP*, pp. VI-3409-3412, Seattle, WA, May 1998.

[9] P. Stoica and R. Moses, *Introduction to Spectral Analysis.* Upper Saddle River, NJ: Prentice Hall, 1997.

[10] Z. Ding, *IEEE Trans. Signal Proc.*, vol. SP-45, pp. 3053-3061, Dec. 1997.

[11] K. Abed-Meraim et al. *IEEE Trans. Signal Processing*, vol. SP-45, pp. 42-55, Jan. 1997.

[12] R.W. Liu and G. Dong, *IEEE Trans. Circuits & Systems-I: Fundamental Theory & Applications*, vol. 44, May 1997.

[13] J.F. Cardoso and A. Souloumiac, *IEE Proc.-F, Radar & Signal Proc.*, vol. 140, pp. 362-370, Dec. 1993.

[14] K. Abed-Meraim et al. *IEEE Trans. Signal Processing*, vol. SP-45, pp. 694-705, March 1997.

[15] D.G. Luenberger, *Optimization by Vector Space Methods.* New York: John Wiley, 1969.

[16] 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.