Proceedings of the 39th IEEE
Conference on Decision and Control
Sydney, Australia • December, 2000

# PROPERTIES OF MATRIX POLYNOMIALS AND MIMO CHANNEL IDENTIFIABILITY FROM SECOND ORDER STATISTICS

*Zhi Ding*
*Department of Electrical and Computer Engineering*
*University of Iowa*
*Iowa City, IA 52242*
*USA*
*zding@eng.uiowa.edu*

*Li Qiu*
*Department of Electrical and Electronic Engineering*
*Hong Kong University of Science and Technology*
*Clear Water Bay, Kowloon*
*Hong Kong*
*eeqiu@ee.ust.hk*

**Abstract**

*Blind channel identification and equalization have attracted much attention recently. Their potential application in wireless communication systems has been explored. For multiple user systems, several algorithms known to date require the assumption that the linear system has a full rank convolution matrix. In fact, this assumption requires that system be irreducible and column reduced. In this paper, we show that some of these algorithms can remain effective even when the rank condition is relaxed. Several algorithms, including the outer-product decomposition, can still be applied with care for irreducible (but not column reduced) multiple-input-multiple-output (MIMO) linear systems.*

## 1. LINEAR MULTI-USER SYSTEM MODELS

### 1.1. Multi-User Data Communication Systems

A synchronous multi-user data communication system can be described using a baseband representation. Assume that the $N_u$ user channels are all linear and causal with impulse response $\{h_u(t), \quad u = 1, 2, \dots N_u\}$, the received output signal can be written as

$$x(t) = \sum_{u=1}^{N_u} \sum_{k=-\infty}^{\infty} s_{k,u} h_u(t - kT) + w(t), \quad s_{k,u} \in \mathcal{A}_u, \quad (1.1)$$

where $T$ is the symbol baud period and $\mathcal{A}_u$ is the input signal set of user $u$. The channel input sequences $\{s_{k,u}\}$ are typically independent for different users and are also i.i.d. The noise $w(t)$ is stationary, white, and independent of channel input sequences $s_{k,u}$, but not necessarily Gaussian.

Note that $h_u(t)$ is a "composite" channel impulse response that includes transmitter and receiver filters as well as the physical channel response. In a typical multi-user system, multiple channels of observations will be available from multiple sensors. If $K_r$ sub-channels (sensors or antennas) exist, then $x(t)$, $h_u(t)$, and $w(t)$ are all $K_r \times 1$ vectors.

### 1.2. Channel Diversity from Oversampling

It has been shown by Tong, et al. [1] that blind channel identification benefits from oversampling the channel outputs. This essentially arises from the spectral diversity available when the channel has bandwidth higher than $1/2T$.

Let the sampling interval be $\Delta = T/p$ where $p$ is an integer. The oversampled discrete signals and responses are

$$x_i \triangleq x(i\Delta), \quad h_u[i] \triangleq h_u(i\Delta) \quad \text{and} \quad w_i \triangleq w(i\Delta). \quad (1.2)$$

The channel output samples are hence

$$x_n = \sum_{u=1}^{N_u} \sum_{k=-\infty}^{\infty} s_{k,u} h_u[n - kp] + w_n.$$

Suppose $\{h_u(t)\}$ has joint finite support $[0, T_h)$, which spans $L+1$ integer baud periods. Define the following notations

$$\begin{aligned}
s_k &\triangleq [s_{k,1} \ s_{k,2} \ \dots \ s_{k,N_u}]'; \\
\mathbf{s}[k] &\triangleq [s'_k \ s'_{k-1} \ \dots \ s'_{k-L-M+1}]'; \\
\mathbf{w}[k] &\triangleq [w_{kp} \ w_{kp+1} \ \dots \ w_{kp-Mp+1}]'; \\
\mathbf{h}_u[i] &\triangleq \begin{bmatrix} h_u[ip] \\ h_u[ip+1] \\ \vdots \\ h_u[ip+p-1] \end{bmatrix}; \\
\mathbf{H}_i &\triangleq [\mathbf{h}_1[i] \ \mathbf{h}_2[i] \ \dots \ \mathbf{h}_{N_u}[i]].
\end{aligned}$$

It is then evident that

$$X_n \triangleq \begin{bmatrix} x_{np} \\ x_{np+1} \\ \vdots \\ x_{np+p-1} \end{bmatrix} = \sum_{i=0}^{L} \mathbf{H}_i s_{n-i}.$$

Now form an $MpK_r \times (L+M)N_u$ block Toeplitz matrix

$$\mathcal{H} = \begin{bmatrix} \mathbf{H}_0 & \mathbf{H}_1 & \dots & \mathbf{H}_L & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_0 & \mathbf{H}_1 & \dots & \mathbf{H}_L & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{H}_0 & \mathbf{H}_1 & \dots & \mathbf{H}_L \end{bmatrix}. \quad (1.3)$$

Clearly, $L$ is the order of the $N_u$ dynamic FIR channels. With these notations, a sampled channel output signal vector of length $Mp$ can be written as

$$\mathbf{x}[k] \triangleq \begin{bmatrix} X_k \\ X_{k-1} \\ \vdots \\ X_{k-M+1} \end{bmatrix} = \mathcal{H}\mathbf{s}[k] + \mathbf{w}[k]. \quad (1.4)$$

### 1.3. Fractional Oversampling

Ordinarily, the oversampling factors are assumed to be an integer. Here, we show that a fractional sampling generates an equivalent multi-user system.

If the oversampling interval is a fraction of $T$, i.e., if $\Delta = qT/p$ ($p, q$ coprime), then the noiseless received signal becomes

$$x_n = \sum_{u=1}^{N_u} \sum_{k=-\infty}^{\infty} s_{k,u} h_u(n\Delta - kT)$$

$$= \sum_{u=1}^{N}\sum_{i=0}^{q-1}\sum_{L=-\infty}^{\infty} s_{Lq+i,u} h_u(n\Delta - LqT - iT)$$
$$= \sum_{u=1}^{N}\sum_{i=0}^{q-1}\sum_{k=-\infty}^{\infty} s_{kq+i,u} h_u(n\Delta - kqT - iT)$$

By defining equivalent signal sequences

$$\bar{s}_{k,i+1+(u-1)q} \triangleq s_{kq+i,u}, \quad i = 0,\ 1,\ \ldots q-1,$$

and equivalent sub-channel responses

$$\bar{h}_{i+1+(u-1)q}[k] = h_u(k\Delta - iT), \quad u = 1,\ \ldots N,\ i = 0,\ \ldots q-1,$$

the received signal can be viewed as an output of $N_u q$ user channels,

$$x_n = \sum_{v=1}^{N_u q}\sum_{k=-\infty}^{\infty} \bar{s}_{k,v}\bar{h}_v[n-kp] + w_n. \qquad (1.5)$$

It is therefore clear that an $N_u$ user system sampled at interval of $\Delta = qT/p$ is equivalent to an $N_u q$ user system. Hence, the blind identification problem of fractionally sampled multi-user system can also be described by equation (1.4) where $\mathcal{H}$ is an $MpK_r \times (M+L)N_u q$ block Toeplitz matrix. This is equivalent to an $N_u p$ input and $K_r p$ output discrete MIMO system.

## 2. MIMO CHANNEL IDENTIFIABILITY AND EQUALIZABILITY

### 2.1. Column-Reduced versus non-Column Reduced Systems

In blind channel identification, the objective is to identify the unknown channel responses $h_u(t)$ based on the channel output $x(t)$ only. The problem of blind identification for such a MIMO dynamic system has been studied in [3, 4, 7, 5, 6, 8] and various algorithms have been proposed.

Define $K = K_r p$ and $N = N_u q$. Then we have a $K \times N$ MIMO transfer function

$$\mathbf{H}(z) = [\sum_{i=0}^{L_1}\mathbf{h}_i^{(1)}z^{-i}\ \ \sum_{i=0}^{L_2}\mathbf{h}_i^{(2)}z^{-i}\ \ \ldots\ \ \sum_{i=0}^{L_N}\mathbf{h}_i^{(N)}z^{-i}]. \quad (2.1)$$

Naturally the maximum channel order is $L = \max L_i$.

It has been established in [4] that one sufficient identification condition for $\mathcal{H}$ to be identifiable from second order statistics is that $\mathcal{H}$ should have full column rank. Equivalently, this blind identifiability condition requires that

- $\mathbf{H}(z) \triangleq \sum_{i=0}^{L}\mathbf{H}_i z^{-i}$ is irreducible, i.e., $\mathbf{H}(z)$ has full column rank for all $z \neq 0$.
- $\mathbf{H}(z)$ is column-reduced [9, 3].

For single input multiple output systems, the second condition is implicit from the first condition. Thus for SIMO systems, $\mathbf{H}(z)$ can be identified using second order statistics directly if it is irreducible. However, for MIMO systems, the additional column-reduced condition does not seem to have any practical physical meaning. Indeed, examples of wireless systems can be provided in which column reduced condition is violated while the system remains irreducible. In this paper, we study whether this condition can be dropped for some of the well established blind identification algorithms.

Our study has its physical significance. First, Tugnait [5] and Meriam [2] have both noted that so long as $\mathbf{H}(z)$ is irreducible, then there exists an FIR equalizer $\mathbf{F}(z)$ for which $\mathbf{F}(z)\mathbf{H}(z) = I$. In other words, the channel is equalizable by FIR equalizers so long as it is irreducible.

The FIR equalizability allows Yule-Walker equations be used to blindly solve for the parameters in $\mathbf{F}(z)$, which can then be used to indirectly identify the equalizer. This is indeed a viable approach and it basically shows that FIR equalizability of $\mathbf{H}(z)$ is equivalent to its blind identifiability by second order statistics. Hence, the requirement that $\mathcal{H}$ be full column rank is not necessary for blind identification using second order statistics. Thus, many algorithms should be modified to accommodate this relaxed non-full rank $\mathcal{H}$ be identified.

Clearly, the necessary dimensional condition for $\mathbf{H}(z)$ to be full rank requires that $K \geq N$. This implication is simple: the number of equivalent multi-channels ($K_r p$) must be no less than the number of equivalent users ($N_u q$).

The critical question that must be answered first is *when it is identifiable*. Here we study how to do this and the mathematical insight behind the blind identification of rank-deficient matrix $\mathcal{H}$.

### 2.2. Necessary and Sufficient Blind Identification Condition

The identification condition hinges on the following theorem.

**Theorem 1** *Let*

$$\mathcal{H} = [\underbrace{\mathcal{H}_1}_{(L+1)N}\ \ \underbrace{\mathcal{H}_2}_{(M-1)N}]. \qquad (2.2)$$

*Assume that* $\mathbf{H}(z)$ *is irreducible and* $M \geq \sum_{i=1}^{N} L_i + \max L_i + 1$. *Then we have*

$$rank(\mathcal{H}) = rank(\mathcal{H}_1) + rank(\mathcal{H}_2).$$

*Moreover,* $\mathcal{H}_1$ *has full rank* $(L+1)N$.

**Proof**: Clearly, this theorem is true when $\mathcal{H}$ has full column rank, i.e., when $\mathbf{H}(z)$ is also column reduced. It can be shown that this is still true when $\mathbf{H}(z)$ is not column-reduced. Details of the proof is given in the appendix.

□

**Theorem 2** *Let* $r_1 = rank(\mathcal{H}_1)$ *and* $r_2 = rank(\mathcal{H}_2)$. *If* $\mathbf{H}(z)$ *is irreducible and* $M \geq 2NL$,

$$\mathcal{H}^H(\mathcal{H}\mathcal{H}^H)^{\#}\mathcal{H} = \begin{bmatrix} \mathbf{I}_{r_1\times r_1} & 0 \\ 0 & B \end{bmatrix} \qquad (2.3)$$

*where* $r_1 = (L+1)N$ *and* $\{.\}^{\#}$ *denotes the pseudo-inverse.*

**Proof**: Define the singular value decompositions of $\mathcal{H}_1$ and $\mathcal{H}_2$ as

$$\begin{aligned} \mathcal{H} &= [\mathcal{H}_1\ \ \mathcal{H}_2] = [U_1\Sigma_1V_1^H\ \ U_2\Sigma_2V_2^H] \\ \text{where}\quad & U_1:\ MK\times r_1,\ \ \Sigma_1:\ r_1\times r_1,\ \ V_1:\ r_1\times r_1, \\ & U_2:\ MK\times r_2,\ \ \Sigma_2:\ r_2\times r_2. \end{aligned}$$

Both $\Sigma_1$ and $\Sigma_2$ are full rank diagonal matrices of singular values.

Since

$$\mathcal{H}\mathcal{H}^H = U\begin{bmatrix} \Sigma_1^2 & 0 \\ 0 & \Sigma_2^2 \end{bmatrix}U^H$$

in which $U = [U_1 \;\; U_2]$ has independent columns according to Theorem 1, $U$ has a unique pseudo-inverse matrix $Q$ such that

$$Q^H U = I.$$

Thus

$$(\mathcal{H}\mathcal{H}^H)^{\#} = Q \begin{bmatrix} \Sigma_1^{-2} & 0 \\ 0 & \Sigma_2^{-2} \end{bmatrix} Q^H,$$

and

$$\mathcal{H}^H(\mathcal{H}\mathcal{H}^H)^{\#}\mathcal{H} = \begin{bmatrix} V_1 & 0 \\ 0 & V_2 \end{bmatrix} \begin{bmatrix} V_1^H & 0 \\ 0 & V_2^H \end{bmatrix}$$

$$= \begin{bmatrix} V_1 V_1^H & 0 \\ 0 & V_2 V_2^H \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{r_1 \times r_1} & 0 \\ 0 & B \end{bmatrix}$$

where $B = V_2 V_2^H$ is identity only if $\mathcal{H}$ has full column rank. □

## 3. OUTER PRODUCT DECOMPOSITION ALGORITHM (OPDA) FOR NON-COLUMN-REDUCED SYSTEMS

The outer-product decomposition algorithm only requires minor modification in the selection of the vector length $M$. We will form an outer-product of the channel parameter matrix

$$\vec{\mathbf{H}} \triangleq \begin{bmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \\ \vdots \\ \mathbf{H}_L \end{bmatrix} \tag{3.1}$$

based on the second order channel output statistics of an MIMO system which is not column-reduced.

First, assume that the channel order $L$ is known. Let

$$H_a \triangleq \begin{bmatrix} \mathbf{H}_0 & \mathbf{H}_1 & \cdots & \mathbf{H}_L & 0\cdots & 0 \\ \mathbf{H}_1 & \mathbf{H}_2 & \cdots & 0 & 0\cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \mathbf{H}_L & 0 & \cdots & 0 & 0\cdots & 0 \end{bmatrix}. \tag{3.2}$$

It can be verified that

$$H_a H_a^H = \tag{3.3}$$

$$\begin{bmatrix} \sum_{k=0}^{L}\mathbf{H}_k\mathbf{H}_k^H & \sum_{k=0}^{L}\mathbf{H}_k\mathbf{H}_{k+1}^H & \cdots & \mathbf{H}_0\mathbf{H}_L \\ \sum_{k=1}^{L}\mathbf{H}_k\mathbf{H}_{k-1}^H & \sum_{k=1}^{L}\mathbf{H}_k\mathbf{H}_k^H & \cdots & \mathbf{H}_1\mathbf{H}_L \\ \vdots & \vdots & \ddots & \\ \sum_{k=L}^{L}\mathbf{H}_k\mathbf{H}_{k-L}^H & \sum_{k=L}^{L}\mathbf{H}_k\mathbf{H}_{k-L+1}^H & \cdots & \mathbf{H}_L\mathbf{H}_L \end{bmatrix}$$

We define $\mathbf{J}$ as a matrix whose first sub-diagonal entries below the main diagonal are unity while all remaining entries are zero. For notational convenience, we define

$$\mathbf{J}^0 = \mathbf{I}, \tag{3.4}$$

$$\mathbf{J}^{-1} = \mathbf{J}'. \tag{3.5}$$

It can be seen that $\mathbf{J}^{-1}\mathbf{A}$ shifts all elements in matrix $\mathbf{A}$ up by one row while $\mathbf{AJ}$ shifts all elements in $\mathbf{A}$ to the left by one column. Thus, we can obtain the outer-product from

$$\Delta \triangleq H_a H_a^H - \mathbf{J}^{-N} H_a H_a^H \mathbf{J}^N$$

$$= \begin{bmatrix} \mathbf{H}_0\mathbf{H}_0^H & \mathbf{H}_0\mathbf{H}_1^H & \cdots & \mathbf{H}_0\mathbf{H}_L^H \\ \mathbf{H}_1\mathbf{H}_0^H & \mathbf{H}_1\mathbf{H}_1^H & \cdots & \mathbf{H}_1\mathbf{H}_L^H \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_L\mathbf{H}_0^H & \mathbf{H}_L\mathbf{H}_1^H & \cdots & \mathbf{H}_L\mathbf{H}_L^H \end{bmatrix} = \vec{\mathbf{H}}\vec{\mathbf{H}}^H$$

Hence, matrix $\Delta$ forms the outer-product of the channel parameter matrix $\vec{\mathbf{H}}$. The singular value decomposition of this outer-product matrix can be used to generate an estimate $\vec{\mathbf{H}}Q$ where $Q$ is an $N \times N$ unitary matrix. Recall from [12] that this memoryless ambiguity is intrinsic to the multi-user blind identification problem and cannot be resolved unless additional information is available and is exploited.

Assume that both the channel input signal and channel noise are white with zero mean. Let their respective covariance matrices be

$$\mathbf{R}_s = E\{\mathbf{s}[k]\mathbf{s}[k]^H\} = \sigma_s^2\mathbf{I}$$

$$\mathbf{R}_w = E\{\mathbf{w}[k]\mathbf{w}[k]^H\} = \sigma_w^2\mathbf{I}.$$

Based on (1.4), the channel output covariance matrix becomes

$$\mathbf{R}_M = E\{\mathbf{x}[k]\mathbf{x}[k]^H\} = \sigma_s^2\mathcal{H}\mathcal{H}^H + \sigma_w^2\mathbf{I} \tag{3.6}$$

Thus, the key step in the algorithm is to find an estimate of the matrix product $H_a H_a^H$ from the statistics of the channel output signal $\mathbf{x}[k]$. Since we focus on the use of second order statistics, our task is to find an estimate of the matrix product $H_a H_a^H$ given the second order statistics of

$$X_k = \sum_{i=0}^{L} \mathbf{H}_i s_{k-i}. \tag{3.7}$$

For notational convenience, define

$$R(n) \triangleq E\{X_k X_{k-n}^H\} = E\{|s_k|^2\}\sum_{i=n}^{L}\mathbf{H}_i\mathbf{H}_{i-n}^H. \tag{3.8}$$

First, it is easy to verify that

$$\mathbf{R}_c = \begin{bmatrix} R(0)-\sigma_w^2 I & R(1) & \cdots & R(L) & \mathbf{0} & \cdots & \mathbf{0} \\ R(1) & R(2) & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ R(L) & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}$$

$$= \sigma_s^2 H_a \mathcal{H}^H.$$

In order to estimate $H_a H_a^H$, recall from Theorem 2 that when $\mathbf{H}(z)$ is irreducible,

$$H_a\mathcal{H}^H(\mathcal{H}\mathcal{H}^H)^{-1}\mathcal{H}H_a^H = H_a \begin{bmatrix} I_{(L+1)N\times(L+1)N} & 0 \\ 0 & B \end{bmatrix} H_a^H$$

$$= H_a H_a^H$$

so long as $M \geq 2NL$.

In other words, we can estimate the channel parameters from

$$\mathbf{R}_c(\mathbf{R}_M - \sigma_w^2\mathbf{I})^{-1}\mathbf{R}_c^H = \sigma_s^2 H_a H_a^H. \tag{3.9}$$

The channel impulse response matrix $\vec{\mathbf{H}}$ can be estimated from the singular value decomposition of the estimate of outer-product matrix $\Delta$.

$$\vec{\mathbf{H}}Q = SVD(\Delta). \tag{3.10}$$

This is the original "outer-product decomposition algorithm" (OPDA) [7] which as also later discovered as the multi-step linear prediction algorithm [8]. Here we show that as opposed to the SSM which requires that the MIMO system be irreducible and column-reduced, the OPDA can identify irreducible MIMO systems so long as $M \geq 2NL$ is chosen.

## 4. MMSE BLIND EQUALIZERS

Assume that $\mathbf{H}(z)$ has full column rank for all $z \neq 0$, $\mathbf{H}(z)$ is thus equalizable and is also identifiable from second order statistics. We now show how an MMSE equalizer can be estimated blindly.

### 4.1. Wiener-Hopf Equation

Assume that the channel input signal $\mathbf{s}[k]$, and channel noise $\mathbf{w}[k]$ are white with zero mean, and the input signal $\mathbf{s}[k]$ has unit variance. The auto-covariance matrices of $\mathbf{s}[k]$ and $\mathbf{w}[k]$ are

$$\mathbf{R}_s(i) = E\{\mathbf{s}[k+i]\mathbf{s}[k]^H\} = \mathbf{J}^{iN}, \quad (4.1)$$

$$\mathbf{R}_w(i) = E\{\mathbf{w}[k+i]\mathbf{w}[k]^H\} = \sigma_w^2 \mathbf{J}^{ip}. \quad (4.2)$$

Denote the FIR equalizer parameter as $\mathbf{g}$ with output signal

$$\mathbf{z}_k = \mathbf{g}^H \mathbf{x}[k].$$

The minimum MSE filter to estimate $\mathbf{s}_{k-i}$ is the solution to the Wiener-Hopf equation,

$$E\{\mathbf{x}[k]\mathbf{x}[k]^H\}\mathbf{g} = E\{\mathbf{x}[k]\mathbf{s}_{k-i}^*\}. \quad (4.3)$$

For white noise and independent input signals, the auto-correlation matrix equals

$$\mathbf{R}_M = \sigma_s^2 \mathcal{H}\mathcal{H}^H + \sigma_w^2 \mathbf{I} = \mathcal{H}\mathcal{H}^H + \sigma_w^2 \mathbf{I} \quad (4.4)$$

while the cross-correlation vector equals

$$E\{\mathbf{x}[k]\mathbf{s}_{k-i}^*\} = \sigma_s^2 \mathbf{H}(i) = \mathbf{H}(i), \quad (4.5)$$

where $\mathbf{H}(i)$ denotes the $(i+1)$-th block column of the channel convolution matrix $\mathcal{H}$.

Given the channel output signal, the auto-correlation matrix can be easily estimated. Thus, the design of MMSE Wiener equalizer depends on our ability to estimate the cross-correlation vector $\mathbf{H}(i)$. We shall now show how this vector can be estimated blindly without training data when $\mathbf{H}(z)$ is irreducible. In our previous work [10, 11], we have shown that if $\mathbf{H}(z)$ is both irreducible and column-reduced, then $\mathcal{H}$ is full rank and equalizers can be blindly estimated. Here our main goal is to show that the removal of the column-reduced condition does not invalidate our blind MMSE equalizer.

### 4.2. Estimation of Cross-Correlation Vector

First, we define that

$$\mathbf{I}_{iN} \triangleq \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{I}_{(M+L-i)N\times(M+L-i)N} \end{bmatrix}_{(M+L)N\times(M+L)N} \quad (4.6)$$

and

$$\Delta\mathbf{I}_{iN} \triangleq \begin{bmatrix} \mathbf{0}_{iN\times iN} & 0 & 0 \\ 0 & \mathbf{I}_{N\times N} & 0 \\ 0 & 0 & 0 \end{bmatrix}_{(M+L)N\times(M+L)N}, \quad (4.7)$$

where $\mathbf{I}_{iN}$ is an identity matrix except for its first $iN$ zero diagonal entries and $\Delta\mathbf{I}_{iN}$ is all zero except for unit entries on its $(iN+1)$ to $(i+1)N$-th diagonal elements. It can be directly shown that $\mathbf{I}_{iN} = \mathbf{J}^{iN}\mathbf{J}^{-iN}$ and $\Delta\mathbf{I}_{iN} = \mathbf{I}_{iN} - \mathbf{I}_{(i+1)N}$.

We exploit second order statistics of the channel output signal contained in its auto-covariance matrices

$$\mathbf{R}_M(i) \triangleq E\{\mathbf{x}[k+i]\mathbf{x}[k]^H\} = \mathcal{H}\mathbf{J}^{iN}\mathcal{H}^H + \sigma_w^2 \mathbf{J}^{iK} \quad (4.8)$$

For simplicity, we first assume that the channel noise is absent such that

$$\mathbf{R}_M(i) = \mathcal{H}\mathbf{J}^{iN}\mathcal{H}^H \quad \text{and} \quad \mathbf{R}_M(0) = \mathcal{H}\mathcal{H}^H. \quad (4.9)$$

Later, we will consider the noisy case when the noise parameter $\sigma_w^2$ must be estimated and subtracted from the covariance matrices.

The critical identity that enables us to estimate the blind MMSE equalizer follows Theorem 2 and (2.3). Denote

$$\begin{aligned} \mathbf{I}_B &\triangleq \mathcal{H}^H(\mathcal{H}\mathcal{H}^H)^{\#}\mathcal{H} \\ &= \begin{bmatrix} \mathbf{I}_{(L+1)N\times(L+1)N} & 0 \\ 0 & B_{(M-1)N\times(M-1)N} \end{bmatrix} \end{aligned}$$

Recall that $\mathbf{J}^i\mathbf{I}_B\mathbf{J}^{-i}$ will shift the elements in $\mathbf{I}_B$ downward by $i$ rows and to the right by $i$-columns. As a result, if $i \geq (M-1)$, then

$$\mathbf{J}^{iN}\mathbf{I}_B\mathbf{J}^{-iN} = \mathbf{I}_{iN}.$$

This equality allows us to use the following method to estimate $\mathbf{H}(i)$ for the MMSE equalizer.

Observe that from the expression of $\mathbf{R}_M(i)$ and (2.3),

$$\begin{aligned} \mathbf{D}_i &\triangleq \mathbf{R}_M(i)\mathbf{R}_M(0)^{\#}\mathbf{R}_M(i)^H \\ &= \mathcal{H}\mathbf{J}^{iN}\mathcal{H}^H(\mathcal{H}\mathcal{H}^H)^{\#}\mathcal{H}\mathbf{J}^{-iN}\mathcal{H}^H \\ &= \mathcal{H}\mathbf{I}_{iN}\mathcal{H}^H, \qquad i \geq (M-1). \end{aligned}$$

Similarly for $i \geq (M-2)$,

$$\mathbf{D}_{i+1} \triangleq \mathbf{R}_M(i+1)\mathbf{R}_M(0)^{\#}\mathbf{R}_M(i+1)^H = \mathcal{H}\mathbf{I}_{(i+1)N}\mathcal{H}^H$$

Then we can form another Hermitian matrix from

$$\begin{aligned} \Delta\mathbf{D}_i \triangleq \mathbf{D}_i - \mathbf{D}_{i+1} &= \mathcal{H}(\mathbf{I}_{iN} - \mathbf{I}_{(i+1)N})\mathcal{H}^H \quad (4.10) \\ &= \mathcal{H}\Delta\mathbf{I}_{iN}\mathcal{H}^H \\ &= \mathbf{H}(i)\mathbf{H}(i)^H, i \geq (M-1). \end{aligned}$$

It is readily seen that matrix $\Delta\mathbf{D}_i$ has rank $N$. An eigen-decomposition will allow us to identify

$$\widehat{\mathbf{H}}(i) = \mathbf{H}(i)\mathbf{Q}, \quad (4.11)$$

in which $\mathbf{Q}$ is an $N \times N$ unitary matrix. As a result, the blind equalizer output in the absence of channel noise is simply

$$\mathbf{z}_k = \mathbf{g}^H\mathbf{x}[k] = \mathbf{Q}^H[\mathbf{R}_M(0)^{-1}\mathbf{H}(i)]^H\mathbf{x}[k] = \mathbf{Q}^H\mathbf{s}_{k-i}^{'},$$

which is a memoryless mixture of $N$ user signals that may need further separation.

A simple generalization can be derived for $i, j \geq (M-1)$ as

$$\begin{aligned} \mathbf{P}_{ij} &\triangleq \mathbf{R}_M(i)\mathbf{R}_M(0)^{\#}\mathbf{R}_M(j)^H \quad (4.12) \\ &\quad -\mathbf{R}_M(i+1)\mathbf{R}_M(0)^{\#}\mathbf{R}_M(j+1)^H \\ &= \mathbf{H}(i)\mathbf{H}(j)^H. \quad (4.13) \end{aligned}$$

## 5. SIMULATIONS

Consider a two user system whose output signal is oversampled by a factor of 3. The two transfer functions are

$$\vec{h}_1(z) = \begin{bmatrix} 0.1 \\ 0.21 \\ 0.07 \end{bmatrix} + \begin{bmatrix} -0.18 \\ -0.23 \\ 0.25 \end{bmatrix} z^{-1} + \begin{bmatrix} 0.4 \\ 0.1 \\ -0.1 \end{bmatrix} z^{-2}$$

$$\vec{h}_2(z) = \begin{bmatrix} .1 \\ -.6 \\ .6 \end{bmatrix} + \begin{bmatrix} .4 \\ .2 \\ .3 \end{bmatrix} z^{-1} + \begin{bmatrix} .1 \\ .1 \\ -.2 \end{bmatrix} z^{-2} + \begin{bmatrix} .8 \\ .3 \\ -.1 \end{bmatrix} z^{-3}$$

The two input signals are both random binary $\pm 1$ with equal probability. The two user signals are scaled so that their output powers are the same. Since $L = 3$ and $N = 2$, we select $M = 2NL = 12$. Both channel lengths are known. By varying the channel output SNR, the normalized MSE between the identified channel responses and the true channel responses are shown in Figure 1. Clearly, the MIMO

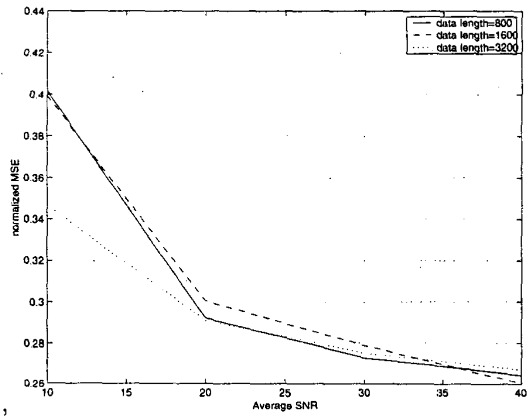


**Figure 1. Mean square identification error.**

system is not column reduced. However, its identification is still feasible, as is shown in this figure and proven in this paper.

## 6. CONCLUSIONS

In this paper, we have shown that several *direct* blind channel identification algorithms based on second order statistics can still be effective for multi-user systems even when $\mathcal{H}$ is not full rank. We show that so long as $\mathbf{H}(z)$ is irreducible (left invertible), it is also identifiable from second order statistics. This means that the condition for perfect (non-blind) equalizability is equivalent to the identification condition based on second order statistics. This important result clarifies the seemingly different identifiability conditions for single user and multiple user systems. It establishes the fact that no additional condition is necessary for the blind identification of unknown linear channels based on second order statistics.

## REFERENCES

[1] L. Tong, G. Xu and T. Kailath, "A New Approach to Blind Identification and Equalization of Multipath Channels," *Proceeding of the 25th Asilomar Conference on Signals, Systems and Computers*, vol.2, pp. 856-860, Pacific Grove. CA. Nov. 1991.

[2] K. Abed-Meriam *et al.*, "Prediction Error Methods for Time-Domain Blind Identification of Multichannel FIR Filters," *Proc. 1995 IEEE ICASSP*, Detroit, MI., May 1995, pp.1968-1971.

[3] D. Slock, "Blind Fractionally-Spaced Equalization, Perfect-Reconstruction Filter Banks and Multichannel Linear Prediction", *Proc. 1994 IEEE ICASSP*, pp. IV:585-588, Adelaide, May 1994.

[4] D. Slock, "Blind Joint Equalization of Multiple Synchronous Mobile Users Using Oversampling and/or Multiple Antennas", *Proc. 1994 Asilomar Conf. on Signals, Systems, and Computers*, pp.:1154-1158, Nov. 1994.

[5] J. K. Tugnait, "Blind Spatio-Temporal Equalization and Impulse Response Estimation for MIMO Channels Using a Godard Cost Function", *IEEE Transactions on Signal Processing*, Sp-45:268-271, Jan. 1997.

[6] K. A. Meriam, P. Loubaton, and E. Moulines, "A Subspace Algorithm for Certain Blind Identification Problems", *IEEE Trans. on Info. Theory*, IT-43:499-512, March 1997.

[7] Z. Ding, "Matrix Outer-Product Decomposition Method for Blind Multiple Channel Identification," *IEEE Trans. Signal Processing*, vol.45, no.12, pp. 3053-3061, Dec. 1997.

[8] D. Gesbert and P. Duhamel, "Robust Blind Identification and Equalization Based on Multi-step Predictors," *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, pp. 3621-3624, München, Germany, May 1997.

[9] T. Kailath, *Linear Systems*, (Section 6.3), *Prentice Hall.* 1980.

[10] J. Shen and Z. Ding, "A Family of Linear Blind Multiple Channel Equalization Algorithms Based on Second Order Statistics," *Proc. 1998 Symposium on Image, Speech, Signal Processing, and Robotics*, II:49-54, Hong Kong, September 1998 (ISBN 962-85361-2-5).

[11] J. Shen and Z. Ding, "Direct Blind MMSE Channel Equalization Based on Second Order Statistics," submitted to *1999 Int. Conf. on Acoustics, Speech, and Signal Proc.*, Pheonix, AZ, 1999.

[12] J. Cardoso, "Source Separation using High Order Moments", in *Proc. IEEE ICASSP*, pp. 2109-22112, 1989.

[13] R. R. Bitmead, S.-Y. Kung, B. D. O. Anderson and T. Kailath, "Greatest Common Divisors via Generalized Sylvester and Bezout Matrices", *IEEE Transactions on Automatic Control*, AC-23:1043-1047, Dec. 1978.

[14] S.-Y. Kung, T. Kailath, and M. Morf, "A Generalized Resultant Matrix for Polynomial Matrices", in *Proc. IEEE Conf. on Decision and Control*, pp.892-895, 1976.

## APPENDIX

Before proving Theorem 1, we first present two lemmae.

**Lemma 1** *[9] For irreducible $K \times N$ polynomial matrix $\mathbf{H}(z)$, there exists a $K \times N$ matrix polynomial $\mathbf{F}(z) = \sum_{i=0}^{m-1} \mathbf{F}_i z^{-i}$ such that $\mathbf{F}(z)^T \mathbf{H}(z) = I_{N\times N}$ Furthermore, the minimum degree of $\mathbf{F}(z)$ is bounded by*

$$m - 1 \geq \sum_{i=1}^{N} L_i.$$

Since the first part of this lemma is simply the well-known Bezout Identity, its proof requires on the construction of $\mathbf{F}(z)$ and its order. Length restriction forces us to omit the proof.

Another important lemma can be found in [14] and [13]

**Lemma 2** *Let $\mathbf{H}(z)$ be an irreducible $K \times N$ ($K > N$) polynomial matrix. Let $\{\bar{L}_i\}$ be the minimal degrees of a minimal polynomial basis for the subspace spanned by columns of $\mathbf{H}(z)$. Then for $M \geq \sum_{i=1}^{N} \bar{L}_i$,*

$$rank[Ker^l(\mathcal{H})] = M(K-N) - \sum_{i=1}^{N} \bar{L}_i$$

*and*

$$rank(\mathcal{H}) = MN + \sum_{i=1}^{N} \bar{L}_i.$$

Now we present the proof of Theorem 1.

**Proof**:

Let $K = K_r p$ and $N = N_u q$. $\mathcal{H}$ is an $MK \times (M+L)N$ block toeplitz matrix generated from the irreducible $K \times N$ polynomial transfer function $\mathbf{H}(z)$.

We first investigate the rank of matrix $\mathcal{H}_2$.

$$\mathcal{H}_2 = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \cdots & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{H}_L & \mathbf{0} & \ddots & \ddots & \mathbf{0} & \mathbf{0} \\ \mathbf{H}_{L-1} & \mathbf{H}_L & \mathbf{0} & \ddots & \mathbf{0} & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \vdots & \\ \mathbf{H}_0 & \mathbf{H}_1 & \dots & \mathbf{H}_L & \mathbf{0} & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{H}_0 & \dots & \mathbf{H}_{L-1} & \mathbf{H}_L \end{bmatrix}. \quad (6.1)$$

Notice that $\mathcal{H}_2$ is an $MK \times (M-1)N$ matrix. To determine the rank of $\mathcal{H}_2$, it is easier to first find the rank of its left kernel.

We now invoke the important (Bezout Identity) Lemma 1. Since $\mathbf{F}(z)^T\mathbf{H}(z) = I$, $\mathbf{F}(z)$ must have full column rank $N$. If we define a $N \times 1$ vector polynomial $\mathbf{a}(z) = \sum_{i=0}^{L} \mathbf{a}_i z^{-i}$ (where $\mathbf{a}_i$ are $N \times 1$ vectors), then

$$[\mathbf{F}(z)\mathbf{a}(z)]^T H(z) = \sum_{i=0}^{L} \mathbf{a}_i^T z^{-i} = \mathbf{a}(z)^T.$$

Denote $\mathbf{u}(z) = \mathbf{F}(z)\mathbf{a}(z) = \sum_{i=0}^{M-1} z^{-i}$. We can define its corresponding parameter vector

$$\mathbf{u}^T = [\mathbf{u}_0^T \ \mathbf{u}_1^T \ \dots \ \mathbf{u}_{M-1}^T].$$

Then

$$\mathbf{u}(z)^T \bar{H}(z) = \sum_{i=0}^{L} \mathbf{a}_i^T z^{-i} \quad \text{if and only if}$$

$$(a) \quad [\mathbf{u}_0^T \ \mathbf{u}_1^T \ \dots \ \mathbf{u}_{M-1}^T]\mathcal{H}_1 = [\mathbf{a}_0^T \ \mathbf{a}_1^T \ \dots \ \mathbf{a}_L^T]$$

$$(b) \quad [\mathbf{u}_0^T \ \mathbf{u}_1^T \ \dots \ \mathbf{u}_{M-1}^T]\mathcal{H}_2 = 0$$

Based on the degrees of freedom in $\mathbf{a}(z)$, this means that, there exist a matrix $\mathbf{U}$ so that

$$\mathbf{U}^T\mathcal{H} = [\mathbf{U}^T\mathcal{H}_1 \ \ \mathbf{U}\mathcal{H}_2] = [\mathbf{A}^T \ \ \mathbf{0}] \quad (6.2)$$

where $\mathbf{A} = I$ with full rank $(L+1)N$.

Let $\mathrm{Ker}^l(A)$ represent the left kernel of a matrix $A$. The relationship of $\mathcal{H}$ and $\mathcal{H}_2$ implies that

$$\mathrm{Ker}^l(\mathcal{H}_2) \supset \mathrm{Ker}^l(\mathcal{H}).$$

Based on (6.2), we find that

$$\mathrm{Ker}^l(\mathcal{H}_2) \supset \{\mathrm{Ker}^l(\mathcal{H}) \cup \mathbf{U}\}$$

As $\mathbf{U}$ is clearly independent of $\mathrm{Ker}^l(\mathcal{H})$, we have

$$rank(\mathrm{Ker}^l(\mathcal{H}_2)) \geq rank(\mathrm{Ker}^l(\mathcal{H}))$$
$$rank(\mathrm{Ker}^l(\mathcal{H})) + rank(\mathbf{U}) = rank(\mathrm{Ker}^l(\mathcal{H})) + (L+1)N.$$

Invoking Lemma 2, the left kernel of $\mathcal{H}_2$ has rank

$$rank(\mathrm{Ker}^l(\mathcal{H}_2)) \geq M(K-N) - \sum_{i=1}^{N} \bar{L}_i + (L+1)N,$$

which means that

$$rank(\mathcal{H}_2) \leq MK - rank(\mathrm{Ker}^l(\mathcal{H}_2)) = MN + \sum_{i=1}^{N} \bar{L}_i - (L+1)N.$$

But since

$$MN + \sum_{i=1}^{N} \bar{L}_i - (L+1)N = rank(\mathcal{H}) - (L+1)N,$$

$$rank(\mathcal{H}_2) \leq rank(\mathcal{H}) - (L+1)N.$$

On the other hand, because $\mathbf{H}(z)$ is irreducible, $\mathbf{H}_0$ has full rank and thus $\mathcal{H}_1$ also has full column rank $(L+1)N$. Consequently, we have the inequality

$$rank(\mathcal{H}_1) + rank(\mathcal{H}_2) \leq rank(\mathcal{H}).$$

Since $\mathcal{H} = [\mathcal{H}_1 \ \ \mathcal{H}_2]$, this means that

$$rank(\mathcal{H}_1) + rank(\mathcal{H}_2) = rank(\mathcal{H})$$

if $deg\{\mathbf{F}(z)\} = m - 1 \geq \sum_{i=1}^{N} L_i$ (Lemma 1) and if (Lemma 2)

$$M \geq \sum_{i=1}^{N} L_i \geq \sum_{i=1}^{N} \bar{L}_i.$$

However,

$$M - 1 = deg(\mathbf{u}(z)) = deg\{\mathbf{F}(z)\} + deg\{\mathbf{a}(z)\} = (m-1) + L.$$

Thus, the sufficient condition on $M$ is that

$$M = m + L \geq \max\{(\sum_{i=1}^{N} L_i + 1 + L), \sum_{i=1}^{N} L_i\} = \sum_{i=1}^{N} L_i + \max L_i + 1$$

$\square$