

# Blind Channel Estimation and Equalization for Multiple FIR Channels

**Shiann-Jeng Yu**

*National Center for High Performance Computing, Hsin-Shi Tainan Country 744, Taiwan  
Email: sjyu@nchc.org.tw*

**Fang-Biau Ueng**

*Department of Electrical Engineering, National Chung Hsing University, 250 Kuo-Kuang Road,  
Taichung 402, Taiwan  
Email: fbueng@ee.nchu.edu.tw*

*Received 15 February 2004; Revised 1 June 2004; Recommended for Publication by Xiaodong Wang*

This paper deals with the problem of blind equalizations based on effective channel order determination for multiple FIR channels. Most popular order determination methods use the eigenvalue decomposition (EVD) technique with an overmodeled data correlation matrix. However, performing the EVD consumes huge computation resources. In this paper, we consider the channel with infinite small leading and tailing terms which is natural for measured microwave radio channels, and develop a computationally simple method for effective channel order determination. Based on multiple-shift property of a data correlation matrix, a new performance index is analyzed. The channel order is determined if the performance index is greater than a threshold. To select the threshold, we model the performance index as an  $F$ -distributed random variable. For a specified confidence level, the threshold can be found from the table. This proposed method does not require EVD, the computation load is much lower than that of the EVD-based methods.

**Keywords and phrases:** second-order statistics, channel equalization, channel order estimation,  $F$ -distribution.

## 1. INTRODUCTION

Blind adaptive equalization of multiple FIR channels without training data available was studied intensively in the literature. Several algorithms have been developed by using the second-order statistics (SOS) [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]. For the SOS-based algorithms, channel order is a crucial parameter for computing the equalization parameters. However, methods for estimating the channel order are quite limited. The most popular methods for order determination, for example, AIC and MDL, are developed based on the information theoretical criteria [11]. However, researches show the AIC and MDL in the measured microwave radio channels are very sensitive to variations in the signal to noise ratio (SNR) and the number of data samples [12]. This prohibits their application for channel order estimation. Due to the nature of practical microwave radio channels having long small leading and tailing channel terms, it has been shown that blind channel equalization algorithms should attempt to model only the significant part of the channel composed of the large impulse response terms [13]. Small leading and tailing terms being modeled in the

blind equalization algorithms in general lead to poor performance and should be avoided [13]. The number of significant part of the channel is referred to the “effective” channel order.

Using numerical analysis arguments and concept of angles of analysis between subspaces and invariant subspace perturbation results, a detection formula [12] is developed to provide a maximally stable decomposition (MSD) of the range space of an overmodeled data correlation into signal subspace and noise subspace. The MSD method compares two consecutive eigenvalues of a correlation matrix for order determination. Existence of a gap between two consecutive eigenvalues makes the channel order being determined. Simulations show that this method provides robustness to variations in the SNR and the number of samples. It also provides information to classify the channels into stable and unstable cases. Like the AIC and MDL methods, the MSD requires the eigenvalues which are obtained from the eigenvalue decomposition (EVD) of an overmodeled correlation matrix. For channels with long small leading and tailing terms, the dimension of the correlation matrix is big in order to ensure enough degrees of

freedom for eigenvalue comparison. The drawback of using big dimension is that the required number of samples will be big too for statistical convergence of the data correlation matrix. Most EVD-based algorithms also suffer from huge computational complexity and numerical sensitivity due to the big dimension. Without the EVD, Gerstacker and Taylor [14] developed a detection algorithm based on the examination of an indicator function constructed from initial channel estimates containing an additional common polynomial factor. However, its performance depends on the accuracy of the channel estimation algorithm.

In this paper, we develop a new detection method without using the EVD. The proposed method is developed based on using multiple-shift property of the data correlation matrix. In the blind channel equalization, the successive data vectors are not statistically independent, thus the data correlation matrix built from the vectors exhibits "shift property." The shift property was first used in the SOS-based equalizer proposed by Tong et al. [1]. They proved that the channel parameters can be estimated from the correlation matrix and a single-shift correlation matrix if the channel order is known. Recently, multiple shift correlation information has been noticed in blind equalization algorithms [8].

We analyze a channel model with infinite small leading and tailing terms and compute two multiple-shift correlation matrices with the shift delay index equal to and greater than the channel order. On the consideration of the worst case of the small channel parameters, we find that trace of the multiple-shift correlation matrix and its complex conjugation can be approximately expressed in the worst case as a comparison of the first and last terms of the significant part of the channel with the small leading and tailing terms. Thus, the performance index can be used as an indicator for the channel order determination. The channel order is determined if the performance index is greater than a threshold. Utilizing the independent and identically distributed (i.i.d.) assumption of the signal and white noise, the performance index is modeled as a random variable with  $F$ -distribution. Proper threshold can be determined from the table by setting a specified confidence level (CL). Unlike the EVD-based methods [11, 12], the proposed method does not restrict the data correlation matrix that should be overmodeled. The required computation load mainly comes from trace operation which is much lower than that of the EVD-based methods.

## 2. PROBLEM FORMULATION

Let  $\mathbf{x}(t) = [x_1(t) \ x_2(t) \ \cdots \ x_p(t)]^T$  be the received data vector of an array with  $p$  sensor elements, where " $T$ " represents the transpose. For digitized data format,  $\mathbf{x}(t)$  is sampled with the sampling rate which is not less than the symbol rate. Consider the measured microwave radio channels, there are long small leading or tailing impulse responses. To model the natural channel, this paper considers the channel with infinite long impulse responses. Therefore, the digitized data

vector can be written in matrix form

$$\begin{aligned} \mathbf{x}(n) &= [\cdots \mathbf{e}_q \ \mathbf{e}_{q-1} \ \cdots \ \mathbf{e}_0 \ \mathbf{h}_0 \ \mathbf{h}_1 \ \cdots \ \mathbf{h}_q \ \mathbf{z}_0 \ \mathbf{z}_1 \ \cdots \ \mathbf{z}_q \ \cdots] \\ &\quad \times \mathbf{s}(n) + \mathbf{w}(n) \\ &= [\mathbf{E} \ \mathbf{H} \ \mathbf{Z}] \mathbf{s}(n) + \mathbf{w}(n), \end{aligned} \quad (1)$$

where  $\mathbf{x}(n) = \mathbf{x}(t = n\Delta T) = [x_1(n) \ x_2(n) \ \cdots \ x_p(n)]^T$  is a  $p \times 1$  vector and  $\Delta T$  is the sampling interval.  $\mathbf{s}(n) = [\cdots s(n+1) \ s(n) \ s(n-1) \ s(n-2) \ \cdots]^T$  and  $s(n-i)$  is the input signal symbol sequence,  $\mathbf{w}(n) = [w_1(n) \ w_2(n) \ \cdots \ w_p(n)]^T$  represents the additive white Gaussian noise vector with dimension  $p \times 1$ . In (1),  $\mathbf{H} = [\mathbf{h}_0 \ \mathbf{h}_1 \ \cdots \ \mathbf{h}_q]$  with size  $p \times (q+1)$  represents the significant part of the channel,  $\mathbf{E} = [\cdots \mathbf{e}_q \ \mathbf{e}_{q-1} \ \cdots \ \mathbf{e}_0]$  and  $\mathbf{Z} = [\mathbf{z}_0 \ \mathbf{z}_1 \ \cdots \ \mathbf{z}_q \ \cdots]$  represent the small leading and tailing terms, respectively. We assume that  $s(n)$  is an i.i.d. zero-mean Gaussian sequence with  $E\{s(i)s^*(j)\} = \delta(i-j)$ , where  $E\{\cdot\}$  is the expectation operation.  $w_i(n)$  is assumed to be the white noise with mean zero and variance  $\sigma_w^2$ , and  $w_i(n)$  is independent of  $s(n)$ . The channel order is equal to  $q$  if there are  $q+1$  significant terms of  $\mathbf{H}$ . It is noted that the channel order may be greater than the number of channels  $p$ . The problem in this paper is to find a method to detect the channel order  $q$  effectively and efficiently.

Several popular order determination methods have been proposed, such as AIC and MDL [11]. Recently, Liavas et al. developed a new detection method based on MSD of the signal and noise subspaces. These methods are based on EVD technique. In order to obtain enough degrees of freedom for eigenvalue comparison,  $m$  data samples of  $\mathbf{x}(n)$  are stacked to form a  $(pm) \times 1$  vector written by  $\mathbf{y}(n) = [\mathbf{x}^T(n) \ \mathbf{x}^T(n-1) \ \cdots \ \mathbf{x}^T(n-m+1)]^T$ . Using (1),  $\mathbf{y}(n)$  can be expressed as follows:

$$\mathbf{y}(n) = \mathbf{A}(h)\mathbf{s}(n) + \mathbf{v}(n), \quad (2)$$

where

$$\mathbf{A}(h) = \begin{bmatrix} \mathbf{E} & \mathbf{H} & \mathbf{Z} & \cdots & \cdots & \cdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \mathbf{E} & \mathbf{H} & \mathbf{Z} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \cdots & \cdots & \cdots & \mathbf{E} & \mathbf{H} & \mathbf{Z} \end{bmatrix} \quad (3)$$

and  $\mathbf{v}(n) = [\mathbf{w}^T(n) \ \mathbf{w}^T(n-1) \ \cdots \ \mathbf{w}^T(n-m+1)]^T$ . Here, we assume the block Toeplitz matrix  $\mathbf{A}(h)$  is of full rank. Let the ensemble correlation matrix of  $\mathbf{y}(n)$  be denoted as  $\mathbf{R}_y = E\{\mathbf{y}(n)\mathbf{y}^H(n)\}$ . The ensemble correlation matrix can be replaced by the sample averaged correlation matrix given by

$$\hat{\mathbf{R}}_y(N) = \frac{1}{N} \sum_{n=1}^N \mathbf{y}(n)\mathbf{y}^H(n), \quad (4)$$

where  $N$  is the number of samples. It is noted that  $\hat{\mathbf{R}}_y(N)$  is the maximum likelihood (ML) estimate of  $\mathbf{R}_y$  in the presence of  $N$  number of samples [15]. Performing the EVD, we have

$$\hat{\mathbf{R}}_y(N) = \sum_{i=1}^{mp} \lambda_i \mathbf{v}_i \mathbf{v}_i^H, \quad (5)$$

where  $\lambda_i$  is the eigenvalue with its corresponding eigenvector  $\mathbf{v}_i$ . Here,  $\lambda_i \geq \lambda_j$  for  $i < j$ . The detection formulas for the AIC and MDL are given by

$$\text{AIC}(k) = -2N(mp - k)G(k) + 2k(2mp - k),$$

$$\text{MDL}(k) = -N(mp - k)G(k) + \frac{1}{2}k(2mp - k) \log(N), \quad (6)$$

respectively, where

$$G(k) = \left( \frac{1}{mp - k} \sum_{i=k+1}^{mp} \log \lambda_i \right) - \log \left( \frac{1}{mp - k} \sum_{i=k+1}^{mp} \lambda_i \right). \quad (7)$$

The Liavas' MSD method [12] is given by

$$\text{MSD}(k) = \begin{cases} \frac{\lambda_{k+1}}{\lambda_k - 2\lambda_{k+1}}, & \text{if } \lambda_{k+1} \leq \frac{1}{3}\lambda_k, \\ 1, & \text{otherwise.} \end{cases} \quad (8)$$

If  $q' = \min(\text{AIC}(k))$  (or  $\min(\text{MDL}(k))$  or  $\min(\text{MSD}(k))$ ), the channel order is detected by  $q = q' - m$ . The basic principle of the EVD-based methods is to separate the correlation matrix into the signal subspace and noise subspace by checking the magnitude of the eigenvalues. In order to obtain better discrimination between the signal and noise subspaces, the dimension is selected to be big. As a result, these methods require much data samples to construct a converged correlation matrix and consume huge computation resources due to the EVD.

### 3. ANALYSIS OF THE MULTIPLE-SHIFT CORRELATION MATRIX

In this section, we analyze a multiple-shift correlation matrix of a channel model with infinite small leading and tailing terms.

#### 3.1. Multiple-shift correlation

Instead of using stacked data vector  $\mathbf{y}(n)$ , we deal with  $\mathbf{x}(n)$  directly. We consider the shift correlation of the  $\mathbf{x}(n)$  vector defined by  $\mathbf{R}_x(k) = E\{\mathbf{x}(n)\mathbf{x}^H(n+k)\}$ . Using (1) and simple calculation, we have

$$\begin{aligned} \mathbf{R}_x(q+k+1) &= \mathbf{e}_{q+k} \mathbf{h}_0^H + \mathbf{e}_{q+k-1} \mathbf{h}_1^H + \cdots + \mathbf{e}_k \mathbf{h}_q^H \\ &+ \mathbf{h}_0 \mathbf{z}_k^H + \mathbf{h}_1 \mathbf{z}_{k+1}^H + \cdots + \mathbf{h}_q \mathbf{z}_{q+k}^H \\ &+ \epsilon(\mathbf{E}, \mathbf{Z}), \end{aligned} \quad (9)$$

where  $\epsilon(\mathbf{E}, \mathbf{Z})$  comprises both the small leading and tailing terms of the channel. Comparing with the first and second terms,  $\epsilon(\mathbf{E}, \mathbf{Z})$  will be a significantly small value and can be ignored. Therefore,  $\mathbf{R}_x(q+k+1)$  can be reduced to

$$\mathbf{R}_x(q+k+1) \approx \begin{bmatrix} \mathbf{E}_k & \mathbf{H} \end{bmatrix} \begin{bmatrix} \mathbf{H}^H \\ \mathbf{Z}_k^H \end{bmatrix}, \quad (10)$$

where

$$\begin{aligned} \mathbf{E}_k &= \begin{bmatrix} \mathbf{e}_{q+k} & \mathbf{e}_{q+k-1} & \cdots & \mathbf{e}_k \end{bmatrix}, \\ \mathbf{Z}_k &= \begin{bmatrix} \mathbf{z}_k & \mathbf{z}_{k+1} & \cdots & \mathbf{z}_{q+k} \end{bmatrix}. \end{aligned} \quad (11)$$

Both have the same size of  $p \times (q+1)$ . Consider a new parameter defined as

$$F(q+k+1) = \text{trace}(\mathbf{R}_x^H(q+k+1)\mathbf{R}_x(q+k+1)). \quad (12)$$

Using (10),  $F(q+k+1)$  can be written as

$$F(q+k+1) \approx \text{trace} \left( \begin{bmatrix} \mathbf{H} & \mathbf{Z}_k \end{bmatrix} \begin{bmatrix} \mathbf{E}_k^H \\ \mathbf{H}^H \end{bmatrix} \begin{bmatrix} \mathbf{E}_k & \mathbf{H} \end{bmatrix} \begin{bmatrix} \mathbf{H}^H \\ \mathbf{Z}_k^H \end{bmatrix} \right). \quad (13)$$

Using the matrix property that  $\text{trace}(\mathbf{AB}) = \text{trace}(\mathbf{BA})$  for matrices  $\mathbf{A}$  and  $\mathbf{B}$ , and after simple straightforward computations, we have

$$\begin{aligned} F(q+k+1) &\approx \text{trace}(\mathbf{H}^H \mathbf{H} (\mathbf{E}_k^H \mathbf{E}_k + \mathbf{Z}_k^H \mathbf{Z}_k)) \\ &+ \text{trace}(\mathbf{H}^H \mathbf{E}_k \mathbf{H}^H \mathbf{Z}_k + \mathbf{E}_k^H \mathbf{H} \mathbf{Z}_k^H \mathbf{H}), \end{aligned} \quad (14)$$

where

$$\begin{aligned} &\text{trace}(\mathbf{H}^H \mathbf{H} (\mathbf{E}_k^H \mathbf{E}_k + \mathbf{Z}_k^H \mathbf{Z}_k)) \\ &\leq \text{trace}(\mathbf{H}^H \mathbf{H}) (\text{trace}(\mathbf{E}_k^H \mathbf{E}_k) + \text{trace}(\mathbf{Z}_k^H \mathbf{Z}_k)), \\ &\text{trace}(\mathbf{H}^H \mathbf{E}_k \mathbf{H}^H \mathbf{Z}_k + \mathbf{E}_k^H \mathbf{H} \mathbf{Z}_k^H \mathbf{H}) \\ &\leq \text{trace}(\mathbf{H}^H \mathbf{E}_k) \text{trace}(\mathbf{H}^H \mathbf{Z}_k) \\ &+ \text{trace}(\mathbf{E}_k^H \mathbf{H}) \text{trace}(\mathbf{Z}_k^H \mathbf{H}). \end{aligned} \quad (15)$$

It is noted that equality of (15) holds if and only if  $\mathbf{E}_k$  and  $\mathbf{Z}_k$  are scalar multiples of  $\mathbf{H}$  [15], that is,

$$\mathbf{E}_k = \eta_E \mathbf{H}, \quad \mathbf{Z}_k = \eta_Z \mathbf{H}, \quad (16)$$

where  $\eta_E$  and  $\eta_Z$  are two complex constants and can be found as  $0 \leq |\eta_E|, |\eta_Z| \leq 1$  due to small leading and tailing terms with respect to the significant part  $\mathbf{H}$ . Using the same arguments above and employing  $\text{trace}(\mathbf{A}^H \mathbf{B} \mathbf{B}^H \mathbf{A}) \leq \text{trace}(\mathbf{A}^H \mathbf{A}) \text{trace}(\mathbf{B}^H \mathbf{B})$  [15], we have

$$\begin{aligned} &\text{trace}(\mathbf{H}^H \mathbf{H} (\mathbf{E}_k^H \mathbf{E}_k + \mathbf{Z}_k^H \mathbf{Z}_k)) \\ &\leq \text{trace}(\mathbf{H}^H \mathbf{H} \mathbf{H}^H \mathbf{H}) (|\eta_E|^2 + |\eta_Z|^2) \\ &\leq \text{trace}(\mathbf{H}^H \mathbf{H})^2 (|\eta_E|^2 + |\eta_Z|^2), \\ &\text{trace}(\mathbf{H}^H \mathbf{E}_k \mathbf{H}^H \mathbf{Z}_k + \mathbf{E}_k^H \mathbf{H} \mathbf{Z}_k^H \mathbf{H}) \\ &\leq (\eta_E \eta_Z^* + \eta_E^* \eta_Z) \text{trace}(\mathbf{H}^H \mathbf{H})^2. \end{aligned} \quad (17)$$

In (17),

$$\text{trace}(\mathbf{H}^H\mathbf{H}) = \sum_{i=0}^q \|\mathbf{h}_i\|^2 \quad (18)$$

is sum of the square norm of all significant part, where  $\|\mathbf{h}_i\|$  represents 2-norm of  $\mathbf{h}_i$ . Using (17), (14) can be rewritten as

$$F(q+k+1) \leq |\eta_E + \eta_Z|^2 \text{trace}(\mathbf{H}^H\mathbf{H})^2. \quad (19)$$

In general, the parameters of the small leading and tailing terms are unknown. The above analysis shows that  $F(q+k+1)$  reaches maximum value when the small leading coefficient matrix  $\mathbf{E}_k$  and the small tailing coefficient matrix  $\mathbf{Z}_k$  are scalar multiples of  $\mathbf{H}$ . Next, consider the case of  $k = -1$ . Using (1), we have

$$\begin{aligned} \mathbf{R}_x(q) &= \mathbf{h}_0\mathbf{h}_q^H + \mathbf{e}_{q-1}\mathbf{h}_0^H + \mathbf{e}_{q-2}\mathbf{h}_1^H + \cdots + \mathbf{e}_0\mathbf{h}_{q-1}^H \\ &\quad + \mathbf{h}_1\mathbf{z}_0^H + \mathbf{h}_2\mathbf{z}_1^H + \cdots + \mathbf{h}_q\mathbf{z}_{q-1}^H + \epsilon'(\mathbf{E}, \mathbf{Z}), \end{aligned} \quad (20)$$

where  $\epsilon'(\mathbf{E}, \mathbf{Z})$  is a term attributed to both the small leading and tailing terms. It can be found that

$$\begin{aligned} F(q) &= \text{trace}(\mathbf{R}_x(q)\mathbf{R}_x(q)) \\ &\geq \text{trace}(\mathbf{h}_q\mathbf{h}_0^H\mathbf{h}_0\mathbf{h}_q^H) \\ &= \|\mathbf{h}_0\|^2\|\mathbf{h}_q\|^2. \end{aligned} \quad (21)$$

### 3.2. A new performance index for channel order detection

From (9),  $\mathbf{R}_x(q+k+1)$  is mainly composed of the product of the channel significant part and the small leading and tailing terms. As shown in (19),  $F(q+k+1)$  is small and not greater than  $|\eta_E + \eta_Z|^2 \text{trace}(\mathbf{H}^H\mathbf{H})^2$ . On the contrary,  $\mathbf{R}_x(q)$  is mainly composed of a product of the channel significant parts  $\mathbf{h}_0$  and  $\mathbf{h}_q$ . For order determination, the significance of either  $\mathbf{h}_0$  or  $\mathbf{h}_q$  should be compared with the small leading and tailing terms of  $\mathbf{E}_k$  and  $\mathbf{Z}_k$ . From (19) and (21), it is observed that if both norm values of  $\mathbf{h}_0$  and  $\mathbf{h}_q$  are big enough,  $F(q)$  should be significantly larger than  $F(q+k+1)$ . Therefore,

$$\text{MSC}(q, k) = \frac{F(q)}{F(q+k+1)} \quad (22)$$

should be a significantly large value if both  $\mathbf{h}_0$  and  $\mathbf{h}_q$  are truly of the significant parts. Using (19) and (21), we find that

$$\text{MSC}(q, k) \geq \frac{\|\mathbf{h}_0\|^2\|\mathbf{h}_q\|^2}{(|\eta_E + \eta_Z| \text{trace}(\mathbf{H}^H\mathbf{H}))^2}. \quad (23)$$

Let  $|\eta_0|^2 = \|\mathbf{h}_0\|^2 / \text{trace}(\mathbf{H}^H\mathbf{H})$  and  $|\eta_q|^2 = \|\mathbf{h}_q\|^2 / \text{trace}(\mathbf{H}^H\mathbf{H})$ , (23) can be rewritten as

$$\text{MSC}(q, k) \geq \left( \frac{|\eta_0\eta_q|}{|\eta_E + \eta_Z|} \right)^2. \quad (24)$$

Because the distribution of the small leading and tailing terms is in general unknown, we consider the worst situation for analysis. The worst situation is to make  $\text{MSC}(q, k)$  minimal when  $\mathbf{E}_k = \eta_E\mathbf{H}$  and  $\mathbf{Z}_k = \eta_Z\mathbf{H}$ . As shown in the equation above, the  $\text{MSC}(q, k)$  is equal to the comparison ratio of the norms of the  $\mathbf{h}_0$  and  $\mathbf{h}_q$  with the norms of all the related small leading and tailing terms. If  $\mathbf{h}_0$  and  $\mathbf{h}_q$  are big enough compared with  $\mathbf{E}_k$  and  $\mathbf{Z}_k$ ,  $\text{MSC}(q, k)$  should be significantly large. On the contrary, neither  $\mathbf{h}_0$  nor  $\mathbf{h}_q$  is not big enough compared with  $\mathbf{E}_k$  and  $\mathbf{Z}_k$ ,  $\text{MSC}(q, k)$  will not be a large enough value. That implies that either  $\mathbf{h}_0$  or  $\mathbf{h}_q$  or both can not be classified into the significant part of the channel. Thus,  $\text{MSC}(q, k)$  can be used as a performance index for determining the channel order  $q$  if the value of  $\text{MSC}(q, k)$  is greater than a threshold.

In this paper, the small leading and tailing terms of  $\mathbf{E}_k$  and  $\mathbf{Z}_k$  are seen random. The  $F(q+k+1)$  for  $k > 0$  will be a random sequence. Consequently, the variation of  $\text{MSC}(q, k)$  for  $k > 0$  depends on the values of  $\mathbf{E}_k$  and  $\mathbf{Z}_k$  of  $F(q+k+1)$ . To smooth out the influence of the individual terms of  $\mathbf{E}_k$  and  $\mathbf{Z}_k$  and to facilitate the utilization of the information, we modify (22) and define the following performance index for channel order determination:

$$\text{MSC}(q) = \frac{F(q)}{\left( (1/(L-q)) \sum_{k=0}^{L-q-1} F(q+k+1) \right)}, \quad (25)$$

where  $L$  is a sufficiently large value.

There is no rule of thumb for selecting  $L$ . On the consideration of computation load and the performance,  $L$  is suggested to be selected, for example, triple the true channel order or more. In the literature, we observe that the true channel order is almost within 6 or 8 for most cases. Therefore,  $L$  is suggested to be within 24.

If  $\mathbf{h}_0$  or  $\mathbf{h}_q$  are big enough, it can be recognized that  $F(q)$  is the hypothesis termed  $\mathcal{H}_1$  and  $F(q')$  for  $q' > q$  can be seen as another hypothesis termed  $\mathcal{H}_0$ . Therefore, the detection formula is given by

$$\text{MSC}(q) \begin{cases} \geq \eta & \text{for } F(q) \text{ in } \mathcal{H}_1, \\ < \eta & \text{for } F(q) \text{ in } \mathcal{H}_0, \end{cases} \quad (26)$$

where  $\eta$  is a detection threshold.

### 3.3. Threshold selection

The above analysis is based on the ensemble correlation matrix. In practice, the correlation matrix is calculated from finite samples. For finite sample averages,  $F(q+k+1)$  can be expressed as

$$\hat{F}(q+k+1) = \text{trace}(\hat{\mathbf{R}}_x^H(q+k+1)\hat{\mathbf{R}}_x(q+k+1)). \quad (27)$$

The performance index is thus given by

$$\begin{aligned} \text{MSC}(q) &= \frac{\hat{F}(q)}{\left( (1/(L-q)) \sum_{k=0}^{L-q-1} \hat{F}(q+k+1) \right)} \quad \text{for } q = 1, 2, \dots, L-2. \end{aligned} \quad (28)$$

For any  $l$ , let

$$\hat{F}(l) = \text{trace}(\hat{\mathbf{R}}_x^H(l)\hat{\mathbf{R}}_x(l)) = \sum_{i=1}^p \hat{g}_i(i), \quad (29)$$

where  $\hat{g}_i(i) = \hat{\mathbf{P}}_i^H \hat{\mathbf{P}}_i$  with  $\hat{\mathbf{P}}_i$  the  $i$ th column of  $\hat{\mathbf{R}}_x(l)$ . Let  $\mathbf{P}_i$  be the  $i$ th column of  $\mathbf{R}_x(l)$ . It is known from the previous section that  $\hat{\mathbf{R}}_y(N)$  is the ML estimate of  $\mathbf{R}_x$ . Each element of  $\hat{\mathbf{R}}_y(N)$  can be seen as a Gaussian random variable with mean equals to the corresponding element of  $\mathbf{R}_y$  for sufficient large number of  $N$ . Since  $\hat{\mathbf{R}}_x(l)$  is an off-diagonal submatrix of  $\hat{\mathbf{R}}_y(N)$  of (4), it is the ML estimate of the correlation matrix of  $\mathbf{R}_x(l)$ ,  $\hat{\mathbf{P}}_i$  can be seen as the Gaussian random vector with mean  $\mathbf{P}_i$ . From [16], it is known that  $\hat{g}_i(i)$  will be a noncentral  $\chi^2$  random variable. The  $\hat{F}(l)$ , sum of all  $\hat{g}_i(i)$ , has the noncentral  $\chi^2$  distribution. According to the probability theory [17],  $\text{MSC}(q)$  has the  $F$ -distribution with 1 and  $L - q$  degrees of freedom, denoted by  $F(1, L - q)$ . Since  $\hat{F}(l)$  for  $l > q$  is of the hypothesis  $\mathcal{H}_0$ , we have confidence that  $0 \leq \text{MSC}(l) < \eta$  at a specified CL for  $l > q$ . It is noted that  $\eta$  is a function of  $l$  and can be written as  $\eta = \eta(L - l)$ . Next, consider the case of  $l = q$ . Since  $\hat{F}(q)$  is the hypothesis  $\mathcal{H}_1$ ,  $\text{MSC}(q)$  should violate the rule of  $\text{MSC}(q) < \eta(L - q)$ . Therefore, the channel order  $q$  can be detected if  $q$  satisfies  $\text{MSC}(q) \geq \eta(L - q)$  and  $\text{MSC}(q + 1) < \eta(L - q - 1)$ .

We summarize the proposed channel order detection procedure as follows.

- Step 1. Specify a CL and find the threshold  $\eta$  from the table.
- Step 2. Compute  $\hat{F}(q)$  by (27) for  $q = 1, 2, \dots, L - 1$ .
- Step 3. Calculate  $\text{MSC}(q)$  by (28) for  $q = 1, 2, \dots, L - 1$ ; set  $q = 1$ .
- Step 4. If  $\text{MSC}(q + 1) < \eta(L - q - 1)$  and  $\text{MSC}(q) \geq \eta(L - q)$ , then go to Step (5), else set  $q = q + 1$  and go back to the comparison of Step (4).
- Step 5. Stop the procedure.

### 3.4. Discussions

The EVD-based methods [11, 12] detect the channel order from the correlation matrix of  $\hat{\mathbf{R}}_y$ , which is with size  $(mp) \times (mp)$ .  $m$  is in general not small to obtain more degrees of freedom for better detection performance [12]. The required computation loads of the EVD-based methods are as high as  $O(m^3 p^3)$ . Instead, the proposed MSC method deals with  $\hat{\mathbf{R}}_x$  directly and the required computation is the trace operation. The computation load is of  $O(Lp^2)$ , where  $L$  can be of the same scale of  $m$ . Therefore, the proposed method consumes much lower computation resources than the EVD-based methods.

Recently, a novel method [14] was proposed without the EVD. The method is based on an indicator function constructed from initial channel estimates. But its performance strongly depends on the accuracy of the channel estimation algorithm. It is noted by [14] that the channel estimation algorithm may yield a highly biased solution that degrades the performance. The method also requires initial guess of the channel order for computing the channel parameters. The

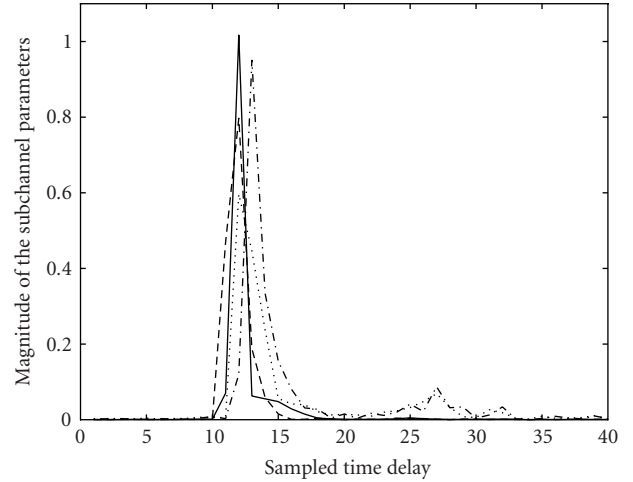


FIGURE 1: The magnitude of the four subchannels.

proposed MSC method detects the channel order from the received data, no channel estimation algorithm is required.

## 4. SIMULATION EXAMPLES

In this section, computer simulations are performed to evaluate the proposed channel order detection method. The AIC, MDL, and MSD are also performed for comparison. We consider the natural channel responses. We have totally 4 subchannels, thus  $p = 4$  in the simulation. Figure 1 shows the magnitude of the four subchannels. It is noted that the channels have long tailing terms. The i.i.d. QPSK signal is used as the input signal.

In the following simulations, we choose  $L = 20$  for the MSC method. The detection thresholds with 90% and 95% confidence levels are used for illustration. The threshold values are listed in Table 1. For the AIC, MDL, and MSD methods, the array multiplicity  $m = 20$  is used, the channel order is determined by  $q = k - m$ , where  $k$  is the number making AIC( $k$ ) or MDL( $k$ ) or MSD( $k$ ) minimal. Figure 2 shows the value of the performance index in comparison with the threshold value. The input SNR is 40 dB and the number of samples used is 500. Here, the SNR is defined in this paper as  $\text{SNR} = 20 \log(1/\sigma_w^2)$ .

The proposed MSC detects the channel order  $q = 2$  by both 90% and 95% confidence levels. In this example, the AIC and MDL detect the channel order  $q = 23$  and 21, respectively, which are much overestimated. The MSD detects  $q = 1$ . Figure 3 shows the output signal to interference plus noise ratio (SINR) of the Tong subspace equalization algorithm at different number of channel order. The results are averaged by 100 independent runs. It is shown that the subspace equalizer reaches the best performance as the channel order is  $q = 2$ . If the channel order is overestimated, the performance gradually degrades. The Tong subspace equalization algorithm could still obtain a certain output SINR level in a certain range of channel order. For example, the averaged output SINR has about 15 dB even the estimated



TABLE 1: The threshold for the MSC.

$L - q$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
90% CL	39.90	8.53	5.54	4.54	4.06	3.78	3.59	3.46	3.36	3.29	3.23	3.18	3.14	3.10	3.07	3.05	3.03	3.01	2.99	2.97
95% CL	161.4	18.51	10.13	7.71	6.61	5.99	5.59	5.32	5.12	4.96	4.84	4.75	4.67	4.60	4.54	4.49	4.45	4.41	4.38	4.35

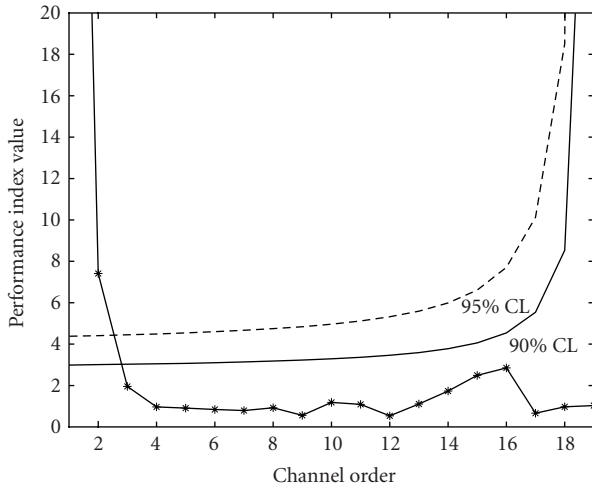


FIGURE 2: The values of the performance index and the threshold under different CL. The solid curve with stars is  $MSC(q)$ , the solid line is the threshold value of 90% CL, and the dashed line is the threshold of 95% CL.

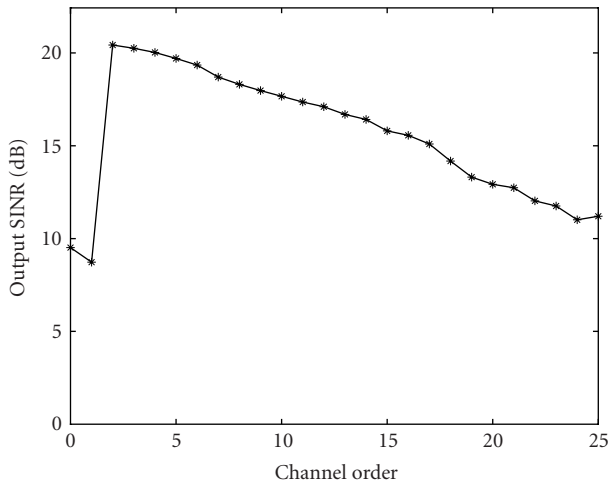


FIGURE 3: The equalization performance of the Tong subspace-based method.

channel order is 17. Constellation plots of the equalization outputs in one of the simulations are shown in Figures 4 and 5 for  $q = 1$  and  $q = 2$ , respectively.

For the sensitivity of the proposed MSC method due to variations of the SNR, we test the proposed MSC method in the input SNR range from 10 dB up to 100 dB, which covers almost all the possible value for successful equalization. The number of independent runs is 100 and the number of samples used is 500. The statistical results are shown in Table 2.

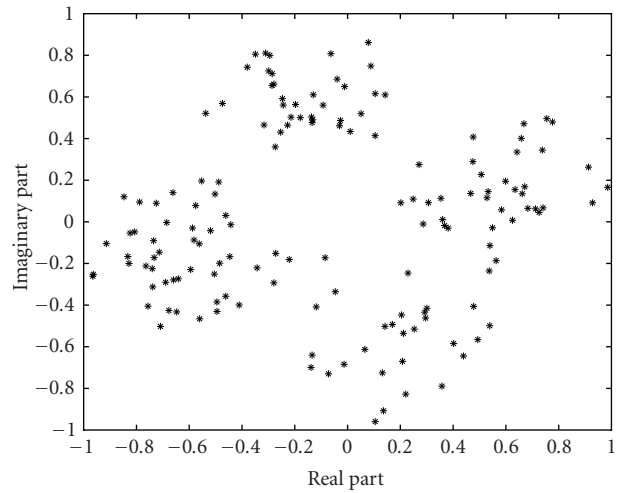


FIGURE 4: Constellation of the output signals using the Tong subspace-based equalizer with estimated channel order  $q = 1$ .

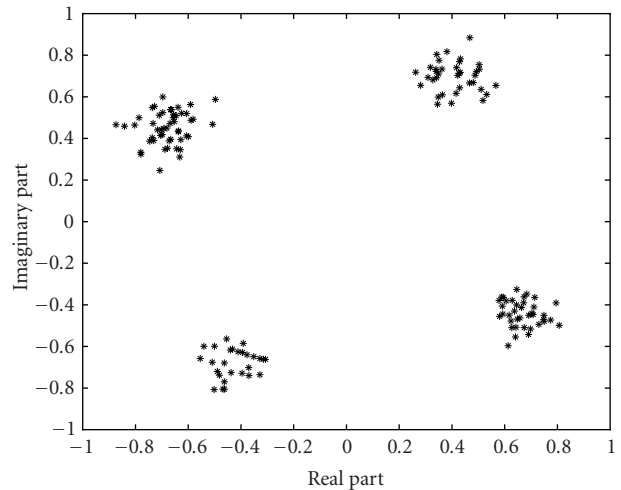


FIGURE 5: Constellation of the output signals using the Tong subspace-based equalizer with estimated channel order  $q = 2$ .

For example, in Table 2(a), 98 runs detect the channel order  $q = 2$  and 2 runs detect the channel order  $q = 3$  with 90% CL in the 100 independent runs. In this example, the MSD detects the channel order  $q = 1$  for all cases. From this table, we find that the MSC method is not sensitive to variations of the input SNR.

To study the effect of a finite number of samples, Table 3 shows the detection results of the MSC method in different number of samples. The number of independent runs is 100

TABLE 2: The statistical results of the proposed MSC method with different values of SNR.

SNR (dB)	$q$	1	2	3	4
10	90% CL	0	98	2	0
	95% CL	0	100	0	0
50	90% CL	0	90	10	0
	95% CL	0	100	0	0
100	90% CL	0	88	12	0
	95% CL	0	99	1	0

(a) The statistical results of the proposed MSC method with SNR = 10 dB.

TABLE 3: The statistical results of the proposed MSC method with different  $N$  values.

$N$	$q$	1	2	3	4
300	90% CL	0	98	2	0
	95% CL	14	86	0	0
500	90% CL	0	88	12	0
	95% CL	0	100	0	0
1000	90% CL	0	54	46	0
	95% CL	0	97	3	0

and the input SNR is 40 dB. In this example, the MSD is quite stable and detects  $q = 1$  for all cases. It is found that the proposed MSC method has the trend to overestimate the channel order, that is,  $q = 3$ . However, from Figure 3,  $q = 3$  is still a reasonably good estimate of the channel order for the Tong subspace equalization algorithm.

## 5. CONCLUSION

An effective channel order detection method for blind channel equalization has been presented. The order detection method use the multiple-shift correlation property of the data. A new performance index has been provided for better understanding of channel order determination. The channel order is detected if the performance index is greater than a threshold. Based on the i.i.d. assumption of the signal and noise, an  $F$ -distribution-based hypothesis testing criterion is used for threshold selection. Unlike the EVD-based methods, the proposed method does not require the EVD which consumes huge computation load for a big-dimension correlation matrix. Simulations show that the proposed method is not sensitive to variation of the input SNR in the test range of 10 dB up to 100 dB and could obtain stable detection probability if the number of samples is not less than 300 for the proposed method with CL.

## REFERENCES

[1] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: a time domain approach," *IEEE Transactions on Information Theory*, vol. 40, no. 2, pp. 340–349, 1994.

- [2] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, no. 2, pp. 516–525, 1995.
- [3] H. Liu and G. Xu, "Closed-form blind symbol estimation in digital communications," *IEEE Trans. Signal Processing*, vol. 43, no. 11, pp. 2714–2723, 1995.
- [4] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Trans. Signal Processing*, vol. 43, no. 12, pp. 2982–2993, 1995.
- [5] Y. Hua, "Fast maximum likelihood for blind identification of multiple FIR channels," *IEEE Trans. Signal Processing*, vol. 44, no. 3, pp. 661–672, 1996.
- [6] M. Tsatsanis and Z. Xu, "Constrained optimization methods for direct blind equalization," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 3, pp. 424–433, 1999.
- [7] J. Mannerkoski and D. P. Taylor, "Blind equalization using least-squares lattice prediction," *IEEE Trans. Signal Processing*, vol. 47, no. 3, pp. 630–640, 1999.
- [8] D. Gesbert, P. Duhamel, and S. Mayrargue, "On-line blind multichannel equalization based on mutually referenced filters," *IEEE Trans. Signal Processing*, vol. 45, no. 9, pp. 2307–2317, 1997.
- [9] L. Tong and Q. Zhao, "Joint order detection and blind channel estimation by least squares smoothing," *IEEE Trans. Signal Processing*, vol. 47, no. 9, pp. 2345–2355, 1999.
- [10] Q. Zhao and L. Tong, "Adaptive blind channel estimation by least squares smoothing," *IEEE Trans. Signal Processing*, vol. 47, no. 11, pp. 3000–3012, 1999.
- [11] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 33, no. 2, pp. 387–392, 1985.
- [12] A. P. Liavas, P. A. Regalia, and J.-P. Delmas, "Blind channel approximation: effective channel order determination," *IEEE Trans. Signal Processing*, vol. 47, no. 12, pp. 3336–3344, 1999.
- [13] A. P. Liavas, P. A. Regalia, and J.-P. Delmas, "Robustness of least-squares and subspace methods for blind channel identification/equalization with respect to effective channel undermodeling/overmodeling," *IEEE Trans. Signal Processing*, vol. 47, no. 6, pp. 1636–1645, 1999.
- [14] W. H. Gerstacker and D. P. Taylor, "Blind channel order estimation based on second-order statistics," *IEEE Signal Processing Letters*, vol. 10, no. 2, pp. 39–42, 2003.
- [15] R. Monzingo and T. Miller, *Introduction to Adaptive Arrays*, John Wiley & Sons, New York, NY, USA, 1980.
- [16] D. Dlugos and R. Scholtz, "Acquisition of spread spectrum signals by an adaptive array," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 37, no. 8, pp. 1253–1270, 1989.
- [17] P. Hoel, S. Port, and C. Stone, *Introduction to Probability Theory*, Houghton Mifflin Company, Boston, Mass, USA, 1971.

**Shiann-Jeng Yu** received the Ph.D. degree from the National Taiwan University in electrical engineering in 1995. From October 1995 to December 2001, he was with National Space Program Office (NSPO) of Taiwan as an Associate Researcher. From January 2001 to July 2002, he was with the National Science Council (NSC) as a Specialist Secretary of Vice Chairman Office. Since August 2002, he has been with the National Center for High Performance Computing (NCHC). He is now an NCHC Deputy Director at the south region office. His specialties and interests are digital signal processing, wireless communication, satellite communication, and grid computing and applications in e-learning.



**Fang-Biau Ueng** received the Ph.D. degree in electronic engineering from the National Chiao Tung University, Hsinchu, Taiwan, in 1995. From 1996 to 2001, he was with National Space Program Office (NSPO) of Taiwan as an Associate Researcher. From 2001 to 2002, he was with Siemens Telecommunication Systems Limited (STSL), Taipei, Taiwan, where he was involved in the design of mobile communication systems. Since February 2002, he has been with the Department of Electrical Engineering, National Chung Hsing University, Taichung, Taiwan. His areas of research interest are wireless communication and adaptive signal processing.

