LINEAR PREDICTION ERROR METHOD FOR BLIND IDENTIFICATION OF TIME-VARYING CHANNELS: THEORETICAL RESULTS

Jitendra K. Tugnait

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

ABSTRACT

Blind channel estimation for SIMO time-varying channels is considered using only the second-order statistics of the data. The time-varying channel is assumed to be described by a complex exponential basis expansion model (CE-BEM). The linear prediction error method for blind identification of time-invariant channels is extended to time-varying channels represented by a CE-BEM. Sufficient conditions for identifiability are investigated. Cyclostationary nature of the received signal is exploited to consistently estimate the time-varying correlation function of the data from a single observation record. The focus of the paper is on certain theoretical issues.

1. INTRODUCTION

Single-input multiple-output (SIMO) time-invariant FIR (finite impulse response) models of received signals arise in several useful baseband-equivalent digital communications and other applications [1]-[4],[6]. The models and approaches discussed in [1]-[4],[6] are based on the assumption that either the channel is time-invariant, or it varies "slowly" so that adaptive (time-recursive) approaches can track its variations with time. However, if the underlying channel undergoes fast (time-selective) fading, then the proposed system models and the approaches based on them will yield degraded performance. In order to handle such situations, it is desirable to consider approaches based on models that explicitly account for fast fading.

Prior work on blind identification/equalization for fast fading channels is sparse. In [7] a basis function expansion approach has been used to convert a time-varying univariate (single user) channel into a time-invariant SIMO (single-input multiple-output) channel. Standard second-order statistics-based subspace methods (as in [2]) are then exploited in [7] for blind channel estimation. Refs. [8] and [10] use complex exponentials as basis functions. [8] achieves diversity by using an antenna array (not always feasible) and treats the single user's components along each exponential basis as a virtual user. [10] considers transmitter signal "design" for single-user.

Notation: $\rho(\mathbf{A})$ denotes the rank of matrix $\mathbf{A}$. Superscripts H and T denote the complex conjugate transpose and the transpose operations, respectively. $\delta(\tau)$ is the Kronecker delta and I_N is the $N \times N$ identity matrix.

2. MODEL ASSUMPTIONS

Consider a time-varying SIMO (single-input multiple-output) FIR (finite impulse response) linear channel with N outputs. Let $\{s(n)\}$ denote an i.i.d. scalar information sequence which is input to the SIMO time-varying channel with discrete-time impulse response $\{\mathbf{h}(n;l)\}$ (N-vector channel response at time n to a unit input at time $n-l$). Then the channel output vector is given by

$$\mathbf{x}(n) := \sum_{l=0}^{L} \mathbf{h}(n;l)s(n-l). \tag{1}$$

In a complex exponential basis expansion representation [8],[10], it is assumed that

$$\mathbf{h}(n;l) = \sum_{q=0}^{Q} \mathbf{h}_q(l)e^{j\omega_q n} \tag{2}$$

where N-column vectors $\mathbf{h}_q(l)$ (for $q = 0, 1, \cdots, Q$) are time-invariant. Eqn. (2) is a basis expansion of $\mathbf{h}(n;l)$ in the time variable n onto complex exponentials with frequencies $\{\omega_q\}$. The noisy measurements of $\mathbf{x}(n)$ are given by

$$\mathbf{y}(n) = \mathbf{x}(n) + \mathbf{v}(n) \tag{3}$$

Assume the following:

- **(H1)** $N > 1$, and $\{\mathbf{h}(n;l)\}$ satisfies (2) where the frequencies ω_q $(q = 0, 1, \cdots, Q)$ are distinct and known. Moreover, $\omega_0 = 0$ (to represent the time-invariant part).
- **(H2)** Let $h_i(n;l)$ denote the i-th component of $\mathbf{h}(n;l)$, $i = 1, 2, \cdots, N$. Define $H_i(z;n) := \sum_{l=0}^{L} h_i(n;l)z^{-l}$. Assume that for any fixed n, $H_i(z;n)$'s for $i = 1, 2, \cdots, N$ have no common zeros (i.e. $N \times 1$ $\mathbf{H}(z;n)$ with i-th component $H_i(z;n)$ is irreducible for any given n).
- **(H3)** $\{s(n)\}$ is zero-mean, white with $E\{|s(n)|^2\} = 1$.
- **(H4)** $\{\mathbf{v}(n)\}$ is zero-mean, white, uncorrelated with $\{s(n)\}$, with $E\{\mathbf{v}(n+\tau)\mathbf{v}^H(n)\} = \sigma_v^2 I_N \delta(\tau)$.

By (**H2**) and [4], it follows that the $[N(K+1)] \times [K+L+1]$ Sylvester matrix $\mathcal{T}_{K;n}(\mathbf{h})$ associated with time-varying SIMO impulse response $\{\mathbf{h}(n;l)\}$ has full-column rank for any $K \geq L-1$ where

$$\mathcal{T}_{K;n}(\mathbf{h}) := \begin{bmatrix} \mathbf{h}(\text{n;0}) & \cdots & \mathbf{h}(\text{n;L}) & \cdots & 0 \\ \vdots & \ddots & & \ddots & \\ 0 & 0 & \mathbf{h}(\text{n;0}) & \cdots & \mathbf{h}(\text{n;L}) \end{bmatrix}. \tag{4}$$

3. CHANNEL IDENTIFICATION

3.1. FIR Linear Predictor

Consider the Hilbert space $\mathcal{H}$ of square integrable complex random variables on a common probability space endowed with the inner product (for scaler complex random variables x_1 and x_2) $< x_1, x_2 >= E\{x_1 x_2^*\}$ where the superscript $*$ denotes complex conjugation (see [5]). Let $Sp\{x_i \in I\}$ denote the subspace of $\mathcal{H}$ generated by the random variables/vectors in the set $\{x_i \in I\}$. Let $H_k(\mathbf{s})$ denote the subspace generated by the past of $\mathbf{x}$ up to time k

$$H_k(\mathbf{x}) := Sp\{\ x_i(k-m),\ \ i = 1, 2, \cdots, N;\ \ m = 0, 1, \cdots\} \tag{5}$$

where $x_i(k)$ is the i-th element of $\mathbf{x}(k)$. Let $H_{k-1,L}(\mathbf{s})$ denote the subspace spanned by a finite past of $\mathbf{s}$

$$H_{k-1,L}(\mathbf{s}) := Sp\{\ s_i(k-m),\ \ i = 1, \cdots, N;\ m = 1, \cdots, L\}. \tag{6}$$

Let $(\mathbf{s}(k)|H_{k-1}(\mathbf{s}))$ denote the orthogonal projection of $\mathbf{s}(k)$ onto the subspace $H_{k-1}(\mathbf{s})$ [5].

This work was supported by NSF Grant CCR-9803850.

Define the $N(K+1)$-column vector

$$X_n := [\mathbf{x}^T(n),\ \mathbf{x}^T(n-1),\ \cdots\ ,\mathbf{x}^T(n-K)]^T \quad (7)$$

and the $K+L+1$-column vector

$$S_n := [s(n),\ s(n-1),\ \cdots\ ,s(n-K-L)]^T. \quad (8)$$

It then follows from (1), (4), (7) and (8) that

$$X_n = \mathcal{T}_{K;n}(\mathbf{h})S_n. \quad (9)$$

Since $\mathcal{T}_{K;n}(\mathbf{h})$ is full column-rank for all $K \geq L-1$ (by (**H2**)), there exists a $[K+L+1]\times[N(K+1)]$ matrix $\mathcal{U}_{K;n}$ such that $\forall n$,

$$\mathcal{U}_{K;n}\mathcal{T}_{K;n}(\mathbf{h}) = I_{K+L+1}. \quad (10)$$

It therefore follows that

$$X_n \in H_{n,K+L}(s),\ S_n \in H_{n,K}(\mathbf{x}),\ H_{n,K+L}(s) = H_{n,K}(\mathbf{x}). \quad (11)$$

Let us rewrite (1) as

$$\mathbf{x}(n) = \mathbf{e}(n) + \widehat{\mathbf{x}}(n|n-1) \quad (12)$$

where

$$\mathbf{e}(n) := \mathbf{h}(n;0)s(n) \quad (13)$$

and

$$\widehat{\mathbf{x}}(n|n-1) := \sum_{l=1}^{L}\mathbf{h}(n;l)s(n-l). \quad (14)$$

Theorem 1. Under (**H1**)-(**H3**), $\{\mathbf{x}(n)\}$ can be decomposed as in (12)-(14) such that

$$E\{\mathbf{e}(n)\mathbf{x}^H(n-m)\} \ = \ 0 \quad \forall m \geq 1, \quad (15)$$

$$\widehat{\mathbf{x}}(n|n-1) \ = \ (\mathbf{x}(n)|H_{n-1}(\mathbf{x})), \quad (16)$$

$$\widehat{\mathbf{x}}(n|n-1) \ \in \ H_{n-1,K+1}(\mathbf{x})\ \forall\ K \geq L-1, \quad (17)$$

$$\widehat{\mathbf{x}}(n|n-1) \ = \ (\mathbf{x}(n)|H_{n-1,K+1}(\mathbf{x}))\ \forall\ K \geq L-1. \quad (18)$$

The decomposition (12) is unique. □

Proof: Eqn. (15) follows from (1), (13) and assumption (**H3**). It follows from (8) and (14) that

$$\widehat{\mathbf{x}}(n|n-1) \in H_{n-1,L}(s) \subset H_{n-1}(s). \quad (19)$$

By (11) and (19) we have

$$H_{n-1,L}(s) \subset H_{n-1,K+1}(\mathbf{x})\ \forall K \geq L-1. \quad (20)$$

Therefore, by (12), (15), (20) and the orthogonal projection theorem (OPT) [5], it follows that (16) and (18) are true as the "error" $\mathbf{e}(n)$ is orthogonal to the data $\mathbf{x}(n-m)$ $(m \geq 1)$, hence to the subspaces $H_{n-1}(\mathbf{x})$ and $H_{n-1,K+1}(\mathbf{x})$ for any $K \geq L-1$. Uniqueness of the decomposition (12) is a consequence of the OPT [5]. □□□

It follows from Theorem 1 that

$$\widehat{\mathbf{x}}(n|n-1) = \sum_{i=1}^{M}\mathbf{A}_{i;n}\mathbf{x}(n-i) \ \text{ for some } \ M \leq L, \quad (21)$$

and some $N \times N$ matrices $\mathbf{A}_{i;n}$s. Using (12) and (21), we have

$$\mathbf{x}(n) = \sum_{i=1}^{M}\mathbf{A}_{i;n}\mathbf{x}(n-i) + \mathbf{e}(n). \quad (22)$$

By (15) and (22), for $m \geq 1$,

$$E\{\mathbf{x}(n)\mathbf{x}^H(n-m)\} = \sum_{i=1}^{M}\mathbf{A}_{i;n}E\{\mathbf{x}(n-i)\mathbf{x}^H(n-m)\}. \quad (23)$$

By the OPT and (18), it is sufficient to consider (23) for $m = l, l+1, \cdots, M$ in order to solve for $\mathbf{A}_{i;n}$s. Using these values of m in (23) we have

$$\mathbf{R}_{xx}(m;n) = \sum_{i=1}^{M}\mathbf{A}_{i;n}\mathbf{R}_{xx}(m-i;n-i), \quad m = 1,2,\cdots,M, \quad (24)$$

where

$$\mathbf{R}_{xx}(m;n) := E\{\mathbf{x}(n)\mathbf{x}^H(n-m)\}. \quad (25)$$

In a matrix formulation, we may write

$$[\mathbf{A}_{1;n},\ \cdots,\ \mathbf{A}_{M;n}]\mathcal{R}_{xxM}^{(n)} = [\mathbf{R}_{xx}(1;n),\ \cdots,\ \mathbf{R}_{xx}(M;n)] \quad (26)$$

where $\mathcal{R}_{xxM}^{(n)}$ denotes an $[NM]\times[NM]$ matrix with its ij-th block element as $\mathbf{R}_{xx}(j-i;n-i)$. Note that $\mathcal{R}_{ssM}^{(n)}$ is not necessarily full rank, therefore, the coefficients $\mathbf{A}_{i;n}$s are not necessarily unique. A minimum norm solution to (26) may be obtained as

$$[\mathbf{A}_{1;n},\ \cdots,\ \mathbf{A}_{M;n}] = [\mathbf{R}_{xx}(1;n),\ \cdots,\ \mathbf{R}_{xx}(M;n)][\mathcal{R}_{xxM}^{(n)}]^{\#} \quad (27)$$

where the superscript # denotes the pseudoinverse.

3.2. Estimation of Leading Coefficient

It follows from (13) and (22) that

$$\mathbf{e}(n) = \mathbf{h}(n;0)s(n) = \mathbf{x}(n) - \sum_{i=1}^{M}\mathbf{A}_{i;n}\mathbf{x}(n-i). \quad (28)$$

Therefore, we have

$$E\{\mathbf{e}(n)\mathbf{e}^H(n)\} = \mathbf{h}(n;0)\mathbf{h}^H(n;0)$$

$$= E\{[\mathbf{x}(n) - \sum_{i=1}^{M}\mathbf{A}_{i;n}\mathbf{x}(n-i)]\mathbf{e}^H(n)\} = E\{[\mathbf{x}(n)\mathbf{e}^H(n)\}$$

$$= \mathbf{R}_{xx}(0;n) - \sum_{i=1}^{M}\mathbf{A}_{i;n}\mathbf{R}_{xx}(i;n) =: \mathcal{D}_n. \quad (29)$$

That is, $\mathcal{D}_n$ can be obtained via (29) using the coefficients $\mathbf{A}_{i;n}$s and the correlations $\mathbf{R}_{xx}(i;n)$. The matrix $\mathcal{D}_n$ is rank one. Carry out an eigenvalue decomposition (EVD) of $\mathcal{D}_n$: let λ_n denote its nonzero eigenvalue and $\mathbf{p}_{\lambda_n}$ be the corresponding unit norm eigenvector. Then

$$\mathbf{h}(n;0) = \alpha_n\sqrt{\lambda_n}\mathbf{p}_{\lambda_n} \ \text{ for some complex } \ \alpha_n \neq 0; \quad (30)$$

note that $|\alpha_n|^2 = 1$. Since the exact value of M is unknown, we take $M = L$ if L is known. If L is unknown, we take $M = L_u$ where L_u is an upperbound on L.

Given $\mathcal{D}_n$ one can determine $\mathbf{h}(n;0)$ only up to a scalar α_n which may be time-varying. We now investigate how to determine α_n's up to a time-invariant scalar. Using (2) and (30), we have ($\mathbf{p}_n := \sqrt{\lambda_n}\mathbf{p}_{\lambda_n}$ and $l \geq 1$)

$$\mathbf{H}\mathbf{B}_n = [\ \alpha_n I_N \quad \alpha_{n+l}I_N \quad \cdots \quad \alpha_{n+lQ}I_N\ ]\,\mathbf{P}_n \quad (31)$$

where

$$\mathbf{H} := [\ \mathbf{h}_0(0) \quad \mathbf{h}_1(0) \quad \cdots \quad \mathbf{h}_Q(0)\ ], \quad (32)$$

$$\mathbf{B}_n := \begin{bmatrix} e^{j\omega_0 n} & e^{j\omega_0 (n+l)} & \cdots & e^{j\omega_0(n+lQ)} \\ e^{j\omega_1 n} & e^{j\omega_1 (n+l)} & \cdots & e^{j\omega_1(n+lQ)} \\ \vdots & \vdots & \cdots & \vdots \\ e^{j\omega_Q n} & e^{j\omega_Q (n+l)} & \cdots & e^{j\omega_Q(n+lQ)} \end{bmatrix}, \quad (33)$$

$$\mathbf{P}_n := \text{block diag}\{\mathbf{p}_n, \mathbf{p}_{n+l}, \cdots, \mathbf{p}_{n+lQ}\}. \quad (34)$$

Similarly, using $\mathbf{h}(i;0)$ for $i = n+l, n+2l, \cdots, n+l(Q+1)$, we have

$$\mathbf{H}\Sigma\mathbf{B}_n = [\ \alpha_{n+l}I_N \quad \alpha_{n+2l}I_N \quad \cdots \quad \alpha_{n+l(Q+1)}I_N\ ]\,\mathbf{P}_{n+l} \quad (35)$$

where

$$\Sigma := \text{diag}\{e^{j\omega_0 l}, e^{j\omega_1 l}, \cdots, e^{j\omega_Q l}\}. \quad (36)$$

From (31) and (35) we have

$$\mathbf{H} = [\ \alpha_n I_N \quad \alpha_{n+l}I_N \quad \cdots \quad \alpha_{n+lQ}I_N\ ]\,\mathbf{P}_n\mathbf{B}_n^{-1}$$

$$= [\ \alpha_{n+l}I_N \quad \alpha_{n+2l}I_N \quad \cdots \quad \alpha_{n+l(Q+1)}I_N\ ]\,\mathbf{P}_{n+l}\mathbf{B}_n^{-1}\Sigma^{-1}. \quad (37)$$

Because of the inherent scale ambiguity in blind identification, we will set $\alpha_n = 1$, to arbitrarily "resolve" it. Let

$$\mathbf{B}_n^{-1} = \left[\ \mathbf{b}_1^H \quad \mathbf{b}_2^H \quad \cdots \quad \mathbf{b}_{Q+1}^H\ \right]^H \quad (38)$$

where $\mathbf{b}_i$ is $1 \times (Q+1)$ $\forall i$. Then (37) can be simplified to yield

$$[\ \alpha_{n+l}I_N \quad \alpha_{n+2l}I_N \quad \cdots \quad \alpha_{n+l(Q+1)}I_N\ ]\,\mathbf{A} = \mathbf{p}_n\mathbf{b}_1 \quad (39)$$

where $\mathbf{A}$ is $[N(Q+1)] \times [Q+1]$ given by

$$\mathbf{A} := \begin{bmatrix} \mathbf{p}_{n+l}\mathbf{b}_1 \\ \mathbf{p}_{n+2l}\mathbf{b}_2 \\ \vdots \\ \mathbf{p}_{n+lQ}\mathbf{b}_Q \\ \mathbf{p}_{n+l(Q+1)}\mathbf{b}_{Q+1} \end{bmatrix} \Sigma^{-1} - \begin{bmatrix} \mathbf{p}_{n+l}\mathbf{b}_1 \\ \mathbf{p}_{n+2l}\mathbf{b}_2 \\ \vdots \\ \mathbf{p}_{n+lQ}\mathbf{b}_Q \\ 0 \end{bmatrix} \quad (40)$$

$$=: \left[\ \mathbf{A}_1^T \quad \mathbf{A}_2^T \quad \cdots \quad \mathbf{A}_{Q+1}^T\ \right]^T \quad (41)$$

Eqns. (40) and (41) may be rewritten as

$$\widetilde{\mathbf{A}}\,[\ \alpha_{n+l} \quad \alpha_{n+2l} \quad \cdots \quad \alpha_{n+l(Q+1)}\ ]^T = \text{vec}\left(\mathbf{b}_1^T\mathbf{p}_n^T\right) \quad (42)$$

$$\text{where} \quad \widetilde{\mathbf{A}} := \left[\ \text{vec}\left(\mathbf{A}_1^T\right) \quad \cdots \quad \text{vec}\left(\mathbf{A}_{Q+1}^T\right)\ \right]. \quad (43)$$

We show in the Appendix that $\rho(\widetilde{\mathbf{A}}) = Q+1$. Therefore, (42) has a unique solution. Given α_{n+il} $(i = 1, 2, \cdots, Q+1)$, we can obtain $\mathbf{h}_q(0)$, $(q = 0, 1, \cdots, Q)$, via (37). This then allows us to obtain $\mathbf{h}(n;0)$ $\forall n$ via (2).

3.3. Estimation of Noise Variance

If $M > L$, then $\mathbf{A}_{i;n} = 0$ for $i > L$ by virtue of (18).

Lemma 1. Under (**H1**)-(**H3**), $\rho(\mathcal{R}_{xxM}^{(n+1)}) \le NM + 1$ for $M \ge L$ where $\rho(A)$ denotes the rank of A. □

Sketch of proof: It follows from (22) that

$$[\ I_N \quad -\mathbf{A}_{1;n} \quad \cdots \quad -\mathbf{A}_{M;n} \quad 0 \quad \cdots \quad 0\ ]\,\mathcal{R}_{xx(M+1)}^{(n+1)}$$

$$= \left[\ \mathbf{h}(n;0)\mathbf{h}^H(n;0) \quad 0 \quad \cdots \quad 0\ \right]. \quad (44)$$

Apply Sylvester's inequality to (44) to deduce the desired result. □□□

Similar to $\mathcal{R}_{xxM}^{(n)}$ in (26), let $\mathcal{R}_{yyM}^{(n)}$ denote a $[NM] \times [NM]$ matrix with its ij-th block element as $\mathbf{R}_{yy}(j-i; n-i) = E\{\mathbf{y}(n-i)\mathbf{y}^H(n-j)\}$. Define similarly $\mathcal{R}_{vvM}^{(n)}$ pertaining to the additive noise $\mathbf{v}(n)$; note that $\mathcal{R}_{vvM}^{(n)} = \sigma_v^2 I_{MN}$. Carry out an EVD of $\mathcal{R}_{yy(M+1)}^{(n+1)} = \mathcal{R}_{xx(M+1)}^{(n+1)} + \sigma_v^2 I_{(M+1)N}$. Then the smallest $N-1$ eigenvalues of $\mathcal{R}_{yy(M+1)}^{(n+1)}$ (for $M \ge L$) equal σ_v^2 because under (**H1**)-(**H3**), $\rho(\mathcal{R}_{xx(M+1)}^{(n+1)}) \le NM + 1$ whereas under (**H4**), $\rho(\mathcal{R}_{yyL_1}) = NM + N$. Thus a consistent estimate $\widehat{\sigma}_v^2$ of σ_v^2 is obtained by taking it as the average of the smallest $N-1$ eigenvalues of $\widehat{\mathcal{R}}_{yy(M+1)}^{(n+1)}$, the data-based consistent estimate of $\mathcal{R}_{yy(M+1)}^{(n+1)}$. Data-based consistent estimation of time-varying correlation function is discussed in Sec. 3.5.

3.4. Channel Identification

Having obtained $\mathbf{e}(n)$ and $\mathbf{h}(n;0)$ (up to a scale factor) for various n's, how do we calculate $\mathbf{h}(n;m)$ for $m = 1, 2, \cdots$ (up to the same scale factor)? This aspect is discuss in this subsection. From (1) and (13), we have

$$E\{\mathbf{x}(n)\mathbf{e}^H(n-m)\}$$

$$= \sum_{i=0}^{L} \mathbf{h}(n;i)E\{s(n-i)s^*(n-m)\}\mathbf{h}^H(n-m;0)$$

$$= \mathbf{h}(n;m)\mathbf{h}^H(n-m;0). \quad (45)$$

Using (28) on the left-side of (45) it follows that

$$\mathbf{B}(m;n) = \mathbf{h}(n;m)\mathbf{h}^H(n-m;0) \quad \text{where} \quad (46)$$

$$\mathbf{B}(m;n) := \mathbf{R}_{xx}(m;n) - \sum_{i=1}^{M} \mathbf{R}_{xx}(m+i;n)\mathbf{A}_{i;n-m}^H. \quad (47)$$

Hence we have $(m = 1, 2, \cdots)$

$$\mathbf{h}(n;m) = \mathbf{B}(m;n)\mathbf{h}(n-m;0)/\|\mathbf{h}(n-m;0)\|^2. \quad (48)$$

3.5. Correlation Function Estimation

The preceding developments are based on the availability of the time-variant correlation function $\mathbf{R}_{xx}(m;n) = E\{\mathbf{x}(n)\mathbf{x}^H(n-m)\}$ of the noise-free signal $\{\mathbf{x}(n)\}$. In practice we only have noisy data $\{\mathbf{y}(n)\}$. We now discuss how to obtain mean-square (m.s.) as well in probability (i.p.) consistent estimates of $\mathbf{R}_{yy}(m;n) = E\{\mathbf{y}(n)\mathbf{y}^H(n-m)\}$ and of σ_v^2 (noise variance), hence of $\mathbf{R}_{xx}(m;n)$, using the representation (2).

From (1), (2) and (**H3**), it follows that

$$\mathbf{R}_{xx}(m;n) = \sum_{q_1=0}^{Q}\sum_{q_2=0}^{Q}\sum_{l=0}^{L} \mathbf{h}_{q_1}(l-m)\mathbf{h}_{q_2}^H(l)e^{j\omega_{q_2}m}e^{j(\omega_{q_1}-\omega_{q_2})n}$$

$$= \sum_{\beta}\left(\sum_{q_1,q_2:\omega_{q_1}-\omega_{q_2}=\beta}\ \sum_{l=0}^{L} e^{j\omega_{q_2}m}\mathbf{h}_{q_1}(l-m)\mathbf{h}_{q_2}^H(l)\right)e^{j\beta n}$$

$$=: \sum_{\beta} \mathbf{C}_{xx}(m;\beta)e^{j\beta n} \quad (49)$$

where the summation in (49) is over all βs for which $\beta = \omega_{q_1} - \omega_{q_2}$, $(q_1, q_2 \in \{0, 1, \cdots, Q\})$. By (**H4**) and (3),

$$\mathbf{R}_{yy}(m;n) = \mathbf{R}_{xx}(m;n) + \sigma_v^2 I_N \delta(m)$$

$$= \sum_{\beta}\left[\mathbf{C}_{xx}(m;\beta) + \sigma_v^2 I_N\delta(m)\delta(\beta)\right]e^{j\beta n}$$

$$=: \sum_{\beta} \mathbf{C}_{yy}(m;\beta)e^{j\beta n}. \tag{50}$$

It follows from (49) and (50) that $\{\mathbf{x}(n)\}$ and $\{\mathbf{y}(n)\}$ are almost cyclostationary sequences with cycle frequencies βs, $\beta = \omega_{q_1} - \omega_{q_2}$ [11],[12]. It follows from [12] that m.s. and i.p. consistent estimates of $\mathbf{C}_{yy}(m;\beta)$ from the measurements $\{\mathbf{y}(n),\ n = 1, 2, \cdots, T\}$ can be formed as

$$\widehat{\mathbf{C}}_{yy}(m;\beta) := (1/T)\sum_{n=1}^{T} \mathbf{y}(n)\mathbf{y}^H(n-m)e^{-j\beta n}. \tag{51}$$

Therefore, a consistent estimate of $\mathbf{R}_{yy}(m;n)$ is obtained as

$$\widehat{\mathbf{R}}_{yy}(m;n) := \sum_{\beta} \widehat{\mathbf{C}}_{yy}(m;\beta)e^{j\beta n}. \tag{52}$$

A consistent estimate of σ_v^2, denote by $\widehat{\sigma_v^2}$, can be obtained as discussed in Sec. 3.2, after replacing $\mathbf{R}_{yy}(m;n)$ with $\widehat{\mathbf{R}}_{yy}(m;n)$ therein. Hence, m.s. and i.p. consistent estimate of $\mathbf{R}_{xx}(m;n)$, $\widehat{\mathbf{R}}_{xx}(m;n)$, follows as

$$\widehat{\mathbf{R}}_{xx}(m;n) := \widehat{\mathbf{R}}_{yy}(m;n) - \widehat{\sigma_v^2} I_N \delta(m). \tag{53}$$

3.6. Practical Implementation

Given data $\mathbf{y}(n),\ n = 1, 2, \cdots, T$. Given frequencies ω_q, $q = 0, 1, \cdots, Q$, in the representation (2), and given L_u, an upperbound on L in (1). Set $M = L_u$. Form the set of all possible second-order cycle frequencies $\beta = \omega_{q_1} - \omega_{q_2}$, $(q_1, q_2 \in \{0, 1, \cdots, Q\}$. Pick $l \geq 1$ in (31). The following steps are executed to implement a practical algorithm to estimate the channel impulse response at a fixed time instant $n_0 \in \{1, 2, \cdots, T\}$.

1) Estimate $\mathbf{C}_{yy}(m;\beta)$ as $\widehat{\mathbf{C}}_{yy}(m;\beta)$ using (51). Estimate $\mathbf{R}_{yy}(m;n)$ as $\widehat{\mathbf{R}}_{yy}(m;n)$ using (52) for $n = n_0, n_0 - 1, \cdots, n_0 - \max(2M, l(Q+1)+M)$. Obtain the estimate of $\widehat{\sigma_v^2}$ of σ_v^2 by following the method discussed in Sec. 3.2, after replacing $\mathbf{R}_{yy}(m;n)$ with $\widehat{\mathbf{R}}_{yy}(m;n)$ therein. Finally estimate $\mathbf{R}_{xx}(m;n)$ as $\widehat{\mathbf{R}}_{xx}(m;n)$ using (53) for $n = n_0, n_0 - 1, \cdots, n_0 - \max(2M, l(Q+1) + M)$.
2) Solve for $\mathbf{A}_{i;n}$ $(i = 1, 2, \cdots, M)$ using (27) after replacing $\mathbf{R}_{xx}(m;n)$ with $\widehat{\mathbf{R}}_{xx}(m;n)$ therein. This needs to be done for $n = n_0, n_0 - 1, \cdots, n_0 - \max(M, l(Q+1))$.
3) Form an estimate $\widehat{\mathcal{D}}_n$ of $\mathcal{D}_n$ using (29) for $n = n_0 - l(Q+1), n_0 - lQ, \cdots, n_0$. Carry out an EVD of $\widehat{\mathcal{D}}_n$ to obtain $\mathbf{p}_n = \sqrt{\lambda_n}\mathbf{p}_{\lambda_n}$. Estimate α_n's for $n = n_0 - l(Q+1), n_0 - lQ, \cdots, n_0$, using (42). This allows us to first estimate $\mathbf{h}_q(0)$, $(q = 0, 1, \cdots, Q)$, via (37) and then to estimate $\mathbf{h}(n;0)$ $\forall n$ via (2).
4) Finally, $\mathbf{h}(n_0, m)$ is obtained via (47) and (48) $\forall m$.

Note that once we estimate $\mathbf{h}(n, m)$ at $Q+1$ distinct values of n, we can obtain $\mathbf{h}(n, m)$ $\forall n, m$ by first uniquely estimating $\mathbf{h}_q(m)$, $q = 0, 1, \cdots, Q$, via an eqn. similar to (31).

4. APPENDIX

Claim 1. $\rho(\mathbf{A}) = Q + 1$ where $\mathbf{A}$ is defined via (40). □
Proof: Since $\omega_0 = 0$, (40) can be rewritten as

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \vdots & \overline{\mathbf{A}} \\ \cdots & \cdots & \cdots \\ & \mathbf{p}_{n+(Q+1)l}\mathbf{b}_{Q+1}\Sigma^{-1} & \end{bmatrix} \tag{54}$$

where $\mathbf{0}$ is $[NQ] \times 1$, $\overline{\mathbf{A}}$ is $[NQ] \times Q$ given by

$$\overline{\mathbf{A}} := \text{block diag.}\{\mathbf{p}_{n+l}, \mathbf{p}_{n+2l}, \cdots, \mathbf{p}_{n+Ql}\}\mathbf{C}_n(\widetilde{\Sigma} - I_Q), \tag{55}$$

$\mathbf{C}_n$ is $Q \times Q$ obtained from $\mathbf{B}_n^{-1}$ by deleting its first column and last row, and $\widetilde{\Sigma} := \text{diag}\{e^{-j\omega_1 l}, \cdots, e^{-j\omega_Q l}\}$. Using the result 0.8.4 on p. 21 of [9] relating $\det(\mathbf{C}_n)$ to $\det(\mathbf{B}_n)$ and minors of $\mathbf{B}_n$, and noting the fact that $\mathbf{B}_n$ is Vandermonde, we can show that $\rho(\mathbf{C}_n) = Q$. Hence, by Sylvester's inequality, $\rho(\overline{\mathbf{A}}) = Q$; (we are assuming that $e^{-j\omega_i l} \neq 1$ for $i = 1, 2, \cdots, Q$). Since the first column of $\mathbf{p}_{n+(Q+1)l}\mathbf{b}_{Q+1}\Sigma^{-1}$ is not a null vector, it then follows that $\rho(\mathbf{A}) = Q + 1$. □□□

Claim 2. $\rho(\widetilde{\mathbf{A}}) = Q + 1$ where $\widetilde{\mathbf{A}}$ is defined via (40), (41) and (43). □
Proof: It follows from (40)-(41) that the q-th column of $\mathbf{A}_i$ is given by $(e^{-j\omega_q l} - 1 + \delta(q - Q - 1))b_{iq}\mathbf{p}_{n+li}$ where b_{iq} denotes the q-th element of the row vector $\mathbf{b}_i$. Therefore, $\rho(\mathbf{A}_i) = 1$ $\forall i$. Since $\rho(\mathbf{A}) = Q+1$, there exist $Q+1$ linearly independent rows of $\mathbf{A}$, no two of which belong to the same submatrix $\mathbf{A}_i$ (else $\rho(\mathbf{A}_i) > 1$). Moreover, if one of these linearly independent rows, $\mathbf{r}$, belongs to $\mathbf{A}_m$, then all other rows of $\mathbf{A}_m$ lie in the one-dimensional subspace spanned by $\mathbf{r}$. The desired result then follows from the definition of $\widetilde{\mathbf{A}}$. □□□

5. REFERENCES

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

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

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

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

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

[6] J.K. Tugnait, L. Tong and Z. Ding, "Single-user channel estimation and equalization," *IEEE Signal Proc. Magazine*, vol. 17, No. 3, pp. 16-28, May 2000.

[7] M.K. Tsatsanis and G.B. Giannakis, "Subspace methods for blind estimation of time-varying FIR channels," *IEEE Trans. Signal Proc.*, SP-45, pp. 3084-3093, Dec. 1997.

[8] H. Liu and G.B. Giannakis, "Deterministic approaches for blind equalization of time-varying channels with antenna arrays," *IEEE Trans. Signal Proc.*, SP-46, pp. 3003-3013, Nov. 1998.

[9] R.A. Horn and C.R. Johnson, *Matrix Analysis.* Cambridge Univ. Press: Cambridge, UK, 1985.

[10] C. Tepedelenlioglu and G.B. Giannakis, "Transmitter redundancy for blind estimation and equalization of time- and frequency-selective channels," *IEEE Trans. Signal Proc.*, SP-48, pp. 2029-2043, July 2000.

[11] W.A. Gardner, *Statistical Spectral Analysis.* Prentice Hall: Englewood Cliffs, NJ, 1988.

[12] A.V. Dandawate and G.B. Giannakis, "Asymptotic theory of mixed time average and kth-order cyclic-moment and cumulant statistics," *IEEE Trans. Inform. Thy.*, IT-41, pp. 216-232, Jan. 1995.