Transmitter Redundancy for Blind Estimation and Equalization of Time- and Frequency-Selective Channels

Cihan Tepedelenlioğlu and Georgios B. Giannakis, *Fellow, IEEE*

***Abstract*—Joint mitigation of time- and frequency-selective fading is an important and challenging problem in mobile communications. Relying on transmitter-induced redundancy, we propose novel channel estimation and symbol recovery approaches for blind identification and equalization of time- and frequency-selective channels, where the time variation is modeled deterministically by a basis expansion. The resulting statistical algorithm enables the usage of a single antenna, dispenses with channel disparity conditions of existing approaches, and allows channel order overestimation. In addition, new deterministic algorithms for generalized OFDM systems are introduced that produce reliable estimates with few data points at high SNR's. Simulations illustrate the approaches developed.**

***Index Terms*—Fading channels, filterbanks, OFDM, system identification, time-varying systems.**

I. Introduction

BLIND identification and equalization algorithms have received considerable attention in the past few years due to their applications in fields such as mobile communications, underwater acoustic communications, and exploration seismology. Especially in communications applications, blind algorithms effectively save bandwidth because they estimate the channel from output only data and, hence, require no training sequences [7]. It is well known that the cyclostationary or multivariate second-order statistics (SOS)-based methods for blind identification suffer from a certain channel disparity condition that can limit the scope of their usefulness (see, e.g., [6], [22], and references therein). It was shown in [20] that precoding at the input using a filterbank enables identification without any channel disparity condition at the expense of redundancy, which can be made arbitrarily small with increasing block lengths.

Most communication channels do not satisfy the time invariance condition, however. It is argued in [27] that under certain circumstances, constant channel assumption might be far from accurate and that channel estimation accuracy can be significantly degraded by time variations. The channel variation could arise due to the relative motion of the transmitter and the receiver, as well as due to oscillator drifts and phase noise coupled with multipath effects. Such effects call for physically justifiable and parsimonious models for the time variation in communication channels.

Due to the Doppler shifts induced by scatterers with different angles of arrival, each channel tap can be modeled as a superposition of complex exponentials with frequencies that depend on these angles. Since the amplitudes and frequencies of the sinusoids vary more slowly than the channel itself, such channel models are useful not only for identification and equalization purposes [5], [23] but for purposes of adaptive coding and prediction of the fading well ahead of the channel's coherence time as well [4].

Modeling the variation as a superposition of sinusoids offers a basis expansion that leads to finite parameterization of the time variation enabling input/output (I/O) identification and, with sufficient diversity, blind estimation of these parameters. This motivated the work in [25], which, under restrictive assumptions on the bases, exploited the diversity in the variation in order to identify a single input single output (SISO) time-varying (TV) channel. Subsequently, [5] established that usage of multiple antennas yields a multi-input multi-output (MIMO) model, where the number of antennas needed for identifiability is on the order of the number of basis functions, but unlike [25], there were no restrictions on the frequencies of the complex exponential bases.

Unlike [5] and [14], which rely on *output* diversity, this paper addresses estimation and equalization of TV channels by exploiting *input* diversity introduced by adding redundancy at the transmitter. In the first part of this work, we generalize the blind identification scheme in [20] to TV channels whose variation is modeled by complex exponential basis functions. This statistical method, where the channel estimate is based on output statistics, is robust at low signal-to-noise ratios (SNR), uses a single antenna output, and imposes no restrictions on channel zero locations (present in [14]) or the frequencies of the basis functions (present in [25]), and moreover, the estimates remain consistent even when the channel order is overestimated. Unlike [23] and [24], the estimation schemes in this paper rely on output data only and do not use training sequences.

Orthogonal frequency division multiplexing (OFDM) systems turn convolutive time-invariant (TI) channels into multiplicative ones. However, when time selectivity is present,

Manuscript received February 26, 1999; revised February 17, 2000. This work was supported by the National Science Foundation under Grant MIP 9424305 and by ARO Grant DAAG55-98-1-0336. Part of the results in this paper appeared in the Proceedings of the 32nd Asilomar Conference on Signals, Systems and Computers, November 1998 and the Proceedings of the IEEE-SP Workshop on Signal Processing Advances in Wireless Communications, May 1999. The associate editor coordinating the review of this paper and approving it for publication was Prof. A. M. Zoubir.

The authors are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: georgios@ece.umn.edu).

Publisher Item Identifier S 1053-587X(00)04948-5.

by duality, the variation has a convolutive effect on the estimated input sequence, causing what is known as intercarrier or interchannel interference, which in turn results in an error floor that increases with the maximum Doppler frequency [18]. Mitigation of fading effects in OFDM was addressed in [12] and [21], where a blind equalization technique in a flat fading environment was proposed. Assuming access to a training sequence, a minimum mean square error (MMSE) equalizer that is robust to channel variations was designed in [13] under a certain separability condition on the channel transfer function. In [10], an equalization algorithm for TV multipath channels was developed, where the time variation is approximated linearly within the block, and the channel is estimated using a pilot signal. In the second part of this paper, we introduce two different generalized OFDM systems for *blind* equalization of time- *and* frequency-selective channels that yield near-perfect channel estimates at high SNR's where the frequencies of the complex exponential basis are approximated by a fine enough FFT grid. Since the time variation within one block is taken into account, unlike conventional OFDM, the proposed generalized OFDM systems are resilient to time variations. The resulting algorithms are deterministic in the sense that they yield exact estimates in the absence of noise, and since they are data based (as opposed to sample correlation based), they do not suffer from finite data length effects. In adverse SNR conditions, the initial estimates obtained via these deterministic algorithms can be further processed and fed to low-complexity adaptive algorithms for noise averaging and tracking purposes.

All the methods herein rely on input diversity that introduces a certain amount of redundancy to the input symbols, thereby reducing the throughput. This redundancy can be viewed as a form of distributed training for the purpose of mitigating the time-varying intersymbol interference (ISI). With each of the methods, we will also provide expressions for how much bandwidth expansion is required to maintain the same throughput and how the bandwidth expansion can be minimized (at the expense of decoding delay) using large block lengths.

The organization of the paper is as follows. Section II introduces the input diversity through filterbank precoders and the basis expansion channel model that is assumed throughout the paper. Section III formulates and solves the blind channel estimation problem and proves the consistency of the method even when the channel order is overestimated. Section IV introduces the two generalized OFDM approaches. Section V discusses input recovery when the channel or its estimates are available, and finally, Section VI verifies that the proposed methods work, sheds some light on their performance, and provides comparisons with existing methods through simulations.

Calligraphic and bold uppercase or lowercase letters will denote matrices (column vectors). Superscript ^{H} will stand for Hermitian, $^{-H}$ for inverse of the Hermitian, superscript * conjugate, $\star$ convolution, ^{T} transpose, $\dagger$ pseudo-inverse, $\otimes$ Kronecker product, $\mathcal{R}$ range, and $\mathcal{N}$ null space. We will adopt the standard MATLAB notation $\mathbf{A}(i:j,k:l)$ to denote the matrix composed of the i through j rows and k through l columns of the matrix $\mathbf{A}$ and use $\mathbf{u}(i:j)$ to denote the vector consisting of the i through j elements of the vector $\mathbf{u}$. We will use $[\mathbf{A}]_{k,m}$ to denote the (k, m)th entry of a matrix $\mathbf{A}$.

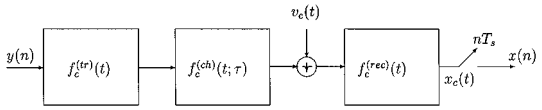

Fig. 1. Continuous-time model for a baseband TV communication system.

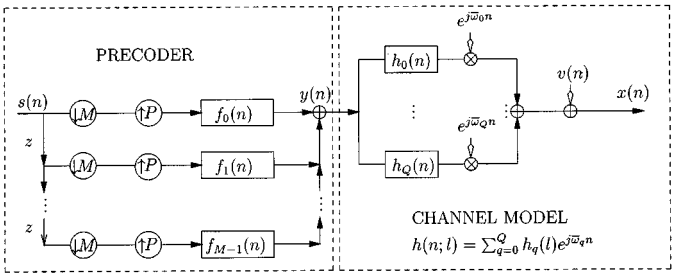

Fig. 2. Precoder and TV channel model.

II. Input Diversity and Channel Model

Consider the communication system depicted in Fig. 1, sampled at the symbol rate $1/T_s$, where subscript $_c$ denotes continuous time. In this section, we will first explain the block transmission scheme that will introduce redundancy to the information sequence $s(n)$, which will aid in blind identification and equalization of the TV channel that the precoded sequence $y(n)$ experiences. Then, we will introduce the channel model where the variation is modeled by a complex exponential basis.

A. Input Diversity

Let the input sequence $s(n)$ be passed through a maximally decimated filterbank, as shown in the precoder part of Fig. 2. The relationship between the information sequence $s(n)$ and the precoded sequence (channel input) $y(n)$ can be expressed as

$$y(n) = \sum_{i=-\infty}^{\infty} \sum_{m=0}^{M-1} s(iM+m) f_m(n-iP). \tag{1}$$

It is well known [26, p. 428] that by blocking $s(n)$ in blocks of size M (the downsampling factor) to get $\mathbf{s}(n) := [s(nM) \cdots s(nM+M-1)]^T$ and, similarly, blocking $y(n)$ in blocks of size P (the upsampling factor), the periodically TV I/O relation in (1) can be represented as a linear MIMO system: $\mathbf{y}(n) := [y(nP) \cdots y(nP+P-1)]^T = \sum_i \mathbf{F}_i \mathbf{s}(n-i)$, where the (k,m)th entry of the $P \times M$ matrix $\mathbf{F}_i$ is given by $[\mathbf{F}_i]_{k,m} = f_{m-1}(iP+k-1)$. Suppose the precoded sequence $y(n)$ is linearly modulated and sent through a linear time-varying, frequency-selective medium (see also Figs. 1 and 2), which, in baseband, discrete-time equivalent form, can be written as

$$x(n) = \sum_{l=0}^{L} h(n;l) y(n-l) + v(n) \tag{2}$$

where $h(n;l) := f_c^{(tr)}(t) \star f_c^{(ch)}(t;\tau) \star f_c^{(rec)}(t)|_{t=nT_s, \tau=lT_s}$. By blocking $x(n)$ into blocks of length P, we obtain the following general I/O relation:

$$\begin{aligned} \mathbf{x}(n) &:= [x(nP) \cdots x(nP+P-1)]^T \\ &= \sum_{j=-\infty}^{\infty} \mathbf{C}_j(n) \sum_{i=-\infty}^{\infty} \mathbf{F}_i \mathbf{s}(n-i-j) + \mathbf{v}(n) \end{aligned} \tag{3}$$

where $[\mathbf{C}_j(n)]_{k,m} = h(nP+k; jP+k-m)$ is a time-varying convolution matrix containing the channel coefficients, and $\mathbf{v}(n) := [v(nP)\cdots v(nP+P-1)]^T$. Let the channel be finite impulse response (FIR) of order L: i) $h(n;l) = 0$, $l \notin [0, L]$, $\forall n$, and the precoder coefficients be chosen such that ii) the length of the precoded block P and the length of the input block M satisfy $M > L$ and $P = M + L$, and iii) $f_m(n) = 0$, $n > P-1$, and $n < 0$ (equivalently, $\mathbf{F}_i = \mathbf{F}_0\delta(i)$). Because the precoder filters $f_m(n)$ are chosen to have length P, only $i = 0$ from (3) will survive, and since $[\mathbf{C}_j(n)]_{k,m} = h(nP+k; jP+k-m)$ and the channel has length L, only $j = 0, 1$ from (3) will provide nonzero contributions. Hence, (3) becomes

$$\mathbf{x}(n) = \mathbf{C}_0(n)\mathbf{F}_0\mathbf{s}(n) + \mathbf{C}_1(n)\mathbf{F}_0\mathbf{s}(n-1) + \mathbf{v}(n). \quad (4)$$

Due to the fact that the channel is FIR of order L, all elements of $\mathbf{C}_1(n)$ except its top right $L \times L$ submatrix are zero. The nonzero elements in $\mathbf{C}_1(n)$ create interblock interference (IBI), i.e., $\mathbf{x}(n)$ depends not only on $\mathbf{s}(n)$ but also on $\mathbf{s}(n-1)$. To eliminate the IBI, one possibility is to set the last L rows of the $(M+L) \times M$ matrix $\mathbf{F}_0$ to zero [the trailing zeros (TZ) approach of [20]], which cancels the term corresponding to $\mathbf{C}_1(n)$ in (4). This means that the MIMO system in (3) has no memory. Defining the $P \times M$ matrix $\mathbf{H}(n) := \mathbf{C}_0^{(:,1:M)}(n)$ to be the first M columns of $\mathbf{C}_0(n)$, and $\mathbf{F} := \mathbf{F}_0(1:M,:)$ the first M rows of $\mathbf{F}_0$, we can write (3) as

$$\underbrace{\mathbf{x}(n)}_{P\times 1} = \underbrace{\mathbf{H}(n)}_{P\times M}\underbrace{\mathbf{F}}_{M\times M}\mathbf{s}(n) + \mathbf{v}(n) \quad (5)$$

where $\mathbf{H}(n)$ has a special banded structure

$$\mathbf{H}(n) := \begin{bmatrix} h(nP;0) & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ h(nP+L;L) & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & h(nP+M-1;0) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \vdots & 0 & h(nP+P-1;L) \end{bmatrix}. \quad (6)$$

Note that setting the last L rows of the $P \times M$ matrix $\mathbf{F}_0$ to zero amounts to creating the sequence $y(n)$ by appending L subsequent zeros after every length-M block $\mathbf{F}\mathbf{s}(n)$. Since we send M precoded information symbols $\mathbf{F}\mathbf{s}(n)$ and L zeros for every block of length $P = M + L$, there is a redundancy of $L/(M+L)$ per block, which we will exploit for channel estimation and symbol recovery.

B. TV Channel Model

The scalar I/O relationship of the TV channel whose input is the output of the precoder (see Fig. 2) is modeled by the following complex exponential basis expansion model (BEM)

$$x(n) = \sum_{q=0}^{Q}\sum_{l=0}^{L_q} h_q(l)e^{j\overline{\omega}_q n}y(n-l) + v(n) \quad (7)$$

where $h_q(l)$ and $\overline{\omega}_q$ are *deterministic* parameters of the channel model. Note that we have allowed the channel order to vary with the basis index q. Let $L := \max_q L_q$ be the order of the TV FIR channel. Then, (7) can be written as

$$x(n) = \sum_{l=0}^{L} \underbrace{\left[\sum_{q=0}^{Q} h_q(l)e^{j\overline{\omega}_q n}\right]}_{h(n;l)} y(n-l) + v(n) \quad (8)$$

which is a basis expansion of $h(n;l)$ in the time variable n onto complex exponentials with frequencies $\{\overline{\omega}_q\}$. Usage of complex exponential bases in capturing the variation in land mobile channels was justified in [5] and [23], where the complex exponential frequencies were associated with each path's Doppler shift.

This model is a deterministic alternative to the statistical modeling of the channel taps and is most useful when a few dominant reflectors are present. Suburban and hilly terrain environments are examples of such a scenario. When the scatterers are dense, a random characterization of the channel might be more appropriate. Jakes [9], [17] has shown that in the presence of an array of radially uniformly spaced independent scatterers with equal magnitudes and uniformly distributed phase angles, the correlation properties of the channel taps (viewed as random processes) are given by

$$\begin{aligned} R_{l_1,l_2}(\tau) &:= E[h(n;l_1)h^*(n+\tau;l_2)] \\ &= J_0(2\pi\tau f_{D_{\max}}/f_s)\delta(l_1 - l_2) \end{aligned}$$

where $J_0(\cdot)$ is the zero-order Bessel function of the first kind, the maximum Doppler shift $f_{D_{\max}}$ is normalized by the sampling frequency f_s, and Krönecker's delta function indicates that different taps are uncorrelated. This so-called uncorrelated scattering constitutes, together with the independence of $R_{l_1,l_2}(\tau)$ from the time index n, the wide sense stationary uncorrelated scattering (WSSUS) fading channel model. The variation of each tap $h(n;l)$ can be simulated by the following sum [9, p. 65]:

$$h(n;l) = \sum_{q=0}^{Q} c_q(l)e^{j\phi_q}\exp[j(2\pi v/\lambda)\cos(2\pi q/Q)n] \quad (9)$$

where

- c_q amplitude of the qth path;
- ϕ_q uniformly distributed random variable in $[0, 2\pi]$;
- λ wavelength corresponding to the carrier frequency;
- v speed of the mobile.

We see that (9) has the same form as the basis expansion in (8) with $h_q(l)$ considered random, having uniformly distributed phases, and $\omega_q = (2\pi v/\lambda)\cos(2\pi q/Q)$, which provides a connection between the BEM and Jakes' model.

Other parametric statistical models of the channel taps include autoregressive moving average models that view each tap as the output of a pole-zero system driven by white noise [2], [24]. The complex exponential BEM that will be assumed throughout this paper can be thought to be the impulse response of a pole-zero system with poles on the unit circle corresponding to frequencies ω_q and with zeros that determine the coefficients $h_q(l)$. Hence, instead of viewing $h(n;l)$ as a random process that is the output of a pole-zero system with white input, as was done in [2] and [24], we effectively view each tap in (8) as the response of a pole-zero system (with poles on the unit circle) to an impulse (see Fig. 3).

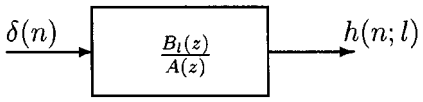

Fig. 3. Channel model.

A related deterministic model was discussed in [21]. The BEM was also motivated and exploited in [4] for adaptive coding and prediction of flat fading communication channels.

Having motivated the model, we would like to write the scalar I/O relation (8) for the $M \times 1$ input $\mathbf{s}(n)$ and the $P \times 1$ output $\mathbf{x}(n)$. Assuming the channel model (8) for the entries in (6), it follows that (5) can be expressed as

$$\mathbf{x}(n) = \sum_{q=0}^{Q} \boldsymbol{\Delta}_q \mathbf{H}_q \mathbf{F} \mathbf{s}(n) e^{j\omega_q n} + \mathbf{v}(n) \tag{10}$$

where $\omega_q := P\overline{\omega}_q$, $\mathbf{H}_q$ is a $P \times M$ Toeplitz matrix given by

$$\mathbf{H}_q := \begin{bmatrix} h_q(0) & 0 & \cdots & 0 \\ \vdots & h_q(0) & \ddots & \vdots \\ h_q(L_q) & \cdots & \ddots & 0 \\ 0 & \ddots & \cdots & h_q(0) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & h_q(L_q) \end{bmatrix} \tag{11}$$

and $\boldsymbol{\Delta}_q := \text{diag}(1, e^{j\overline{\omega}_q}, \cdots, e^{j\overline{\omega}_q(P-1)})$. For brevity, we will denote $\overline{\mathbf{H}}_q := \boldsymbol{\Delta}_q \mathbf{H}_q$ for the rest of the paper.

Estimation of the frequencies $\overline{\omega}_q$, or the frequencies ω_q corresponding to the block model in (10), relies on Fourier analysis of the vector output $\mathbf{x}(n)$'s moments. Since $y(n)$ is not a stationary sequence, we cannot apply the methods in [5] for frequency estimation to (8) directly because in [5], the channel input was stationary. That is why for the case where the input $s(n)$ is precoded with a filterbank, we need to use the stationarized vector model in (10) or, equivalently, the polyphase components of the scalar output sequence $x(nP + k)$ to apply the results in [5] for frequency estimation. Consider the moments of $x(nP+k)$. Notice that one can obtain sums and differences of the frequencies $\omega_0, \cdots, \omega_Q$ from the spectral analysis of the moments of $x(nP + k)$, as long as the corresponding input moments are nonzero. For example

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} E[x(nP+k_1)x^*(nP+k_2)]e^{-j\alpha n}$$

has spectral lines at $\alpha = \omega_{q_1} - \omega_{q_2}, q_1, q_2 = 0, \cdots, Q$, i.e., all possible differences of the frequencies. The corresponding unconjugated correlation would have spectral lines at $\alpha = \omega_{q_1} + \omega_{q_2}, q_1, q_2 = 0, \cdots, Q$ whenever $E[s^2(n)] \neq 0$, but since most constellations satisfy the circular symmetry condition $E[s^2(n)] = 0$, for this class of constellations, we can only get all possible differences of the frequencies using second-order statistics. Since the set $\{\omega_{q_1} - \omega_{q_2}, q_1, q_2 = 0, \cdots, Q\}$ is not sufficient to recover ω_q, this information must be complemented with information one can get from $x(n)$ about ω_q via fourth-order moments. Having to use fourth-order statistics, due to its slow convergence, might seem discouraging, but it is well known that frequency estimation via (even fourth-order) periodogram methods exhibits fast convergence properties [16].

Frequency estimation will also reveal the number of dominant frequencies Q of the model. In [5], we also detail how to estimate Q from the rank properties of a certain matrix constructed from output data.

Note that having to work with the polyphase components $x(nP+k)$, i.e., P-spaced samples of the output sequence $x(n)$ might create aliasing problems if the frequencies to be estimated are too large. Since, as we will see in the next section, choosing P large decreases the redundancy and increases the throughput, it is of interest to see how large the block length P could be chosen without running into aliasing problems in frequency estimation. For a number of practical systems, the maximum Doppler frequency does not exceed $f_{D_{\max}} = 100$ Hz. If the symbol duration is about $T_s = 3.7 \times 10^{-6}$ s, as is the case in the global system for mobile (GSM) standard, then the maximum normalized Doppler frequency $f_{D_{\max}} T_s = 3.7 \times 10^{-4}$, which corresponds to a $\omega_Q = \overline{\omega}_Q P = (2\pi)3.7 \times 10^{-4} P \approx 2.5 \times 10^{-3} P$. Then, in the case that we would need to use fourth-order statistics, the maximum frequency we would have to estimate using the spectral analysis of the output moments would be $4\overline{\omega}_Q P \approx 10^{-2} P$. This would allow P to be on the order of several hundreds for $4\overline{\omega}_Q P < 2\pi$ to hold, which agrees with block lengths seen in practical systems.

More details on frequency estimation from the scalar output $x(n)$ along with further justification for the validity of the basis expansion model are provided in [5] and [23].

III. Statistical Method

In this section, we provide a method that relies on the output second-order statistics to estimate the channel parameters. The parameter estimates are formulated as a function of the output statistics. This provides us with noise averaging capability that also becomes clear in the simulations.

A. Channel Estimation

Given a finite record of the vector output $\mathbf{x}(n)$ in (10), the estimates of frequencies $\{\omega_q\}_{q=0}^{Q}$, (and, hence, of $\{\boldsymbol{\Delta}_q\}_{q=0}^{Q}$), we would like to estimate the matrices $\{\mathbf{H}_q\}_{q=0}^{Q}$ in (10). Let

$$\mathbf{C}_{xx}(\alpha) := \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} E[\mathbf{x}(n)\mathbf{x}^H(n)]e^{-j\alpha n}$$

be the cyclic correlation matrix of the data $\mathbf{x}(n)$ at cycle α. For identifiability, we need to assume the following.

a1) Matrix $\mathbf{R}_{ss} := E[\mathbf{s}(n)\mathbf{s}^H(n)]$ is full rank. Assumption a1) means that the input sequence $s(n)$ is persistently exciting (p.e.) of order at least M, which is required even in I/O identification schemes and is satisfied by most input streams in digital communications. To facilitate clear exposition of the basic idea, we make the following auxiliary (but not necessary) assumption.

a2) The spacings of frequencies $\omega_0, \cdots \omega_Q$ are such that for $\omega_{q_2} \neq \omega_{q_1}, \omega_{q_2} - \omega_{q_1} = \omega_{q_4} - \omega_{q_3} \Rightarrow \omega_{q_4} = \omega_{q_2}$, i.e., all possible differences of the frequencies are distinct.

In Appendix A, we discuss how to dispense with this assumption and yet restore identifiability.

Using (10) and a2), assuming the noise $\mathbf{v}(n)$ is stationary and $\omega_{q_1} \neq \omega_{q_2}$, we obtain

$$\mathbf{C}_{xx}(\omega_{q_1} - \omega_{q_2}) = \boldsymbol{\Delta}_{q_1}\mathbf{H}_{q_1}\mathbf{F}\mathbf{R}_{ss}\mathbf{F}^H\mathbf{H}_{q_2}^H\boldsymbol{\Delta}_{q_2}^H. \tag{12}$$

Since $\{\boldsymbol{\Delta}_q\}_{q=0}^Q$ are invertible, we can write $\boldsymbol{\Delta}_{q_1}^{-1}\mathbf{C}_{xx}(\omega_{q_1} - \omega_{q_2})\boldsymbol{\Delta}_{q_2}^{-H} = \mathbf{H}_{q_1}\mathbf{F}\mathbf{R}_{ss}\mathbf{F}^H\mathbf{H}_{q_2}^H$. Because $\mathbf{R}_{ss}$ is full rank by assumption, $\mathbf{F}$ is full column rank by design, and $\mathbf{H}_{q_2}$ is always full rank due to its structure [see (11)], we infer that $\mathcal{R}(\boldsymbol{\Delta}_{q_1}^{-1}\mathbf{C}_{xx}(\omega_{q_1} - \omega_{q_2})\boldsymbol{\Delta}_{q_2}^{-H}) = \mathcal{R}(\mathbf{H}_{q_1})$. Now, we show how to recover $\mathbf{H}_{q_1}$, having a basis for its range space.

Let $\tilde{\mathbf{U}}$ denote the orthogonal complement of the singular vectors that span $\mathcal{R}(\boldsymbol{\Delta}_{q_1}^{-1}\mathbf{C}_{xx}(\omega_{q_1} - \omega_{q_2})\boldsymbol{\Delta}_{q_2}^{-H})$. Then, subspace orthogonality implies that

$$\tilde{\mathbf{U}}^H\mathbf{H}_{q_1} = \mathbf{0}. \tag{13}$$

Suppose that for every $l = 0, \cdots, L_{q_1}$, the lth column of $\tilde{\mathbf{U}}$, $\tilde{\mathbf{u}}_l = [\tilde{u}_l(0) \cdots \tilde{u}_l(L_{q_1})]^T$ is used to generate the $(L_{q_1}+1) \times M$ Hankel matrix $\mathbf{U}_l$ whose first column is $[\tilde{u}_l(0) \cdots \tilde{u}_l(L_{q_1})]^T$ and whose last row is $[\tilde{u}_l(L_{q_1}) \cdots \tilde{u}_l(P-1)]$, and let $\mathbf{h}_q := [h_q(0) \cdots h_q(L_{q_1})]^T$. Because both

$$\mathbf{h}_{q_1}^H[\mathbf{U}_1 \cdots \mathbf{U}_{L_{q_1}}] = \mathbf{0}^H \tag{14}$$

and (13) yield identical equations for the unknown coefficients in $\mathbf{h}_{q_1}$, we conclude that $[\mathbf{U}_1 \cdots \mathbf{U}_{L_{q_1}}]$ contains $\mathbf{h}_{q_1}^H$ in its left null space. Therefore, the steps involved in estimating $\mathbf{H}_{q_1}$ are as follows.

Step 1) Estimate $\mathbf{C}_{xx}(\omega_{q_1} - \omega_{q_2})$ by

$$\hat{\mathbf{C}}_{xx}(\omega_{q_1} - \omega_{q_2}) = (1/N)\sum_{n=0}^{N-1}\mathbf{x}(n)\mathbf{x}^H(n) \cdot e^{-j(\omega_{q_1}-\omega_{q_2})n}.$$

Step 2) From the singular value decomposition (SVD) of $\boldsymbol{\Delta}_{q_1}^{-1}\hat{\mathbf{C}}_{xx}(\omega_{q_1} - \omega_{q_2})\boldsymbol{\Delta}_{q_2}^{-H}$, compute $\hat{\tilde{\mathbf{U}}}$.

Step 3) Calculate $\hat{\mathbf{h}}_{q_1}$ as the singular vector corresponding to the minimum singular value of $[\hat{\mathbf{U}}_1 \cdots \hat{\mathbf{U}}_{L_{q_1}}]^H$ mentioned in (14).

The steps above outline the estimation of $\mathbf{H}_q$ up to a scale factor. Since the estimation of each $\mathbf{H}_q$ involves a scale ambiguity, we know how to estimate $\beta_q\mathbf{H}_q$, where β_q, $q = 0, \cdots Q$ are unknown scalars. In order to estimate the input, $\{\beta_q\}_{q=0}^Q$ needs to be estimated. Since there is an inherent scalar ambiguity (common to all blind methods) that cannot be resolved, without loss of generality, we will set $\beta_0 = 1$. Substituting the estimates back in (12), we obtain $\mathbf{C}_{xx}(\omega_0 - \omega_q) = \boldsymbol{\Delta}_0\mathbf{H}_0\mathbf{F}\mathbf{R}_{ss}\mathbf{F}^H\beta_q\mathbf{H}_q^H\boldsymbol{\Delta}_q^H$ from which we can estimate β_q for all q. In short, the second-order cyclic statistics $\mathbf{C}_{xx}(\omega_{q_1} - \omega_{q_2})$ contain enough information to reduce the Q scalar ambiguities $\{\beta_q\}_{q=1}^Q$ into one (which is inherent to the problem).

Recovery of $\mathbf{h}_{q_1}$ from (14) can be done uniquely up to a scale factor. This was proved in [20] for matrices with the same structure for the time-invariant channel identification problem, where the Vandermonde or generalized Vandermonde left nullspace structure of the channel matrix was used, but $h_{q_1}(l) \neq 0$ for $l = 0$ and $l = L$ was assumed. In what follows, we remove this restriction by assuming only $h_{q_1}(l) \neq 0$ for some l and provide an alternative proof of the fact that (14) admits a unique solution for $\mathbf{h}_{q_1}$.

Since, as mentioned before, (13) provides the same set of equations as (14) for the unknown coefficients in $\mathbf{h}_{q_1}$, if one admits a unique solution, so does the other. Therefore, it suffices to show that if two matrices $\mathbf{H}_q$ and $\tilde{\mathbf{H}}_q$ with the structure in (11) simultaneously satisfy (13) (i.e., have the same basis for their left nullspace), then they must be proportional. Because $\mathbf{H}_q$ and $\tilde{\mathbf{H}}_q$ are full rank and have the same basis for their left null space, it follows that they have the same range space, i.e., $\mathbf{H}_q\mathbf{T} = \tilde{\mathbf{H}}_q$, where $\mathbf{T}$ is an $M \times M$ invertible matrix. We need to show that $\mathbf{T}$ is a multiple of an identity matrix. Suppose that L_q is the maximum l for which $h_q(l) \neq 0$. This means that $\mathbf{H}(P - L + L_q + 1 : P, :) = \mathbf{0}$, and since $\mathcal{R}(\mathbf{H}_q) = \mathcal{R}(\tilde{\mathbf{H}}_q)$, it follows that L_q is also the maximum l for which $\tilde{h}_q(l) \neq 0$; hence, $h_q(l)$ and $\tilde{h}_q(l)$ have the same length. Now, consider the product $\mathbf{H}(1 : P-L+L_q, :)\mathbf{T}(:, m) = \tilde{\mathbf{H}}(1 : P-L+L_q, m)$. However, this is a restatement of the convolution between the length L_q+1 sequence $\{h_q(l)\}_{l=0}^{L_q}$ and $\{\mathbf{T}(k, m)\}_{k=1}^M$ (the mth column of $\mathbf{T}$), which could be rewritten in the z domain as

$$H_q(z)t_m(z) = \tilde{H}_q(z)z^{m-1}, \qquad m = 1 \cdots M \tag{15}$$

where

$$t_m(z) := \sum_{k=1}^{M}\mathbf{T}(k, m)z^{-(k-1)}, \quad \text{and}$$

$$H_q(z) := \sum_{l=0}^{L_q}h_q(l)z^{-l}.$$

Since $h_q(l)$ and $\tilde{h}_q(l)$ have the same length, $H_q(z)$ and $\tilde{H}_q(z)$ have the same order. It follows from (15) that $t_1(z) = \lambda$, which is a constant. Consequently, $H_q(z)\lambda = \tilde{H}_q(z)$ and $t_m(z) = \lambda z^{m-1}$, which is equivalent to saying that $\mathbf{T} = \lambda\mathbf{I}$.

Because $s(n)$ and, hence, $y(n)$ have finite moments, and the subchannels $h_q(l)$ have finite length, the estimator for the cyclic correlation matrix $\mathbf{C}_{xx}(\alpha)$ in Step 1 is mean-square consistent [3]. Following steps 1–3, it is easy to see that $\mathbf{h}_{q_1}$ is a continuous function of $\mathbf{C}_{xx}(\omega_{q_1} - \omega_{q_2})$; hence, the estimate $\hat{\mathbf{h}}_{q_1}$ is also mean-square consistent.

This blind method for TV channel identification requires only a single antenna for identifiability. The price paid is the throughput reduction since $P = M + L$ symbols are sent for every M information symbols. However, what matters is the ratio M/P, which can be made arbitrarily close to 1 for sufficiently large block lengths.

Note that $h_q(l) \neq 0$ for some l is enough for $\mathbf{H}_q$ to be full rank, which implies that the present input diversity scheme like its TI counterpart in [20] but, unlike many second-order based output diversity methods (e.g., [15]), does not impose any restrictions on the channel coefficients for identifiability.

Another important feature of the input diversity scheme herein is its robustness to channel order overestimation. In the following subsection, we show how using an $\overline{L} > L$ in our input diversity scheme still yields an accurate estimate of the channel coefficients.

B. Robustness to Order Overestimation

In this section, we will show that the proposed input diversity scheme works with an upper bound $\overline{L} > \max_q L_q$ to the channel order. The analysis here pertains to the noise-free case, and to simplify the arguments, a2) will be assumed. In the construction or estimation of a given matrix, to clarify whether the true channel order L, or its upper bound $\overline{L}$, is used, (L) or $(\overline{L})$ will be associated with that matrix for clarity.

Suppose we know an upper bound $\overline{L} > L_q, \forall q$. Then, we will design our transmitter precoder matrix $\mathbf{F}_0(\overline{L})$ to have $\overline{L}$ trailing zeros and choose $P = M + \overline{L}$. This implies that the last $\overline{L} - L_q$ rows of $\boldsymbol{\Delta}_q^{-1}\mathbf{C}_{xx}(\omega_q - \omega_r)\boldsymbol{\Delta}_r^{-H}$ will be zero. In estimating $\mathbf{H}_q(\overline{L})$, one relies on the equation $\hat{\hat{\mathbf{U}}}^H\hat{\mathbf{H}}_q(\overline{L}) = \mathbf{0}$, or equivalently

$$\mathcal{R}(\hat{\mathbf{H}}_q(\overline{L})) = \mathcal{R}\left(\begin{bmatrix} \boldsymbol{\Delta}_q^{-1}(L)\hat{\mathbf{C}}_{xx}^{(L)}(\omega_q - \omega_r)\boldsymbol{\Delta}_r^{-H}(L) \\ \hat{\mathbf{0}}_{(\overline{L}-L_q)\times P} \end{bmatrix}\right) \tag{16}$$

where $\hat{\mathbf{0}}_{(\overline{L}-L_q)\times P}$ is a nonzero matrix that asymptotically goes to $\mathbf{0}$ as the estimation accuracy increases. Any solution that we get from (16) will be of the form

$$\hat{\mathbf{H}}_q(\overline{L}) = \begin{bmatrix} \hat{\mathbf{H}}_q(L) \\ \hat{\mathbf{0}}_{(\overline{L}-L_q)\times M} \end{bmatrix} \tag{17}$$

where $\hat{\mathbf{H}}_q(L)$ is estimated using $\mathcal{R}(\hat{\mathbf{H}}_q(L)) = \mathcal{R}(\boldsymbol{\Delta}_q^{-1}\hat{\mathbf{C}}_{xx}^{(L)}(\omega_q - \omega_r)\boldsymbol{\Delta}_r^{-H})$. However, this would be the precise equation that would be used to estimate $\hat{\mathbf{H}}_q(L)$ if we had the true order L_q. Due to the fact that the method is consistent when we have access to the true channel order, because of (17), the channel estimate with an overestimated order is also consistent. Note the following remarks.

Remark 1: Since the statistical method herein is a generalization of the channel identification method proposed in [20], it follows that the filterbank precoding scheme in [20] that assumes a TI channel is also resilient to channel order overestimation. This fact was observed in [20] in the simulations but was not shown analytically.

Remark 2: Even though we have shown here that usage of any $\overline{L}$ that satisfies $\overline{L} > \max_q L_q$ guarantees identifiability, because the amount of redundancy introduced per block is $\overline{L}$, it is important to choose $\overline{L}$ as close to $\max_q L_q$ as possible in order not to reduce the information rate more than necessary.

The method utilized for channel identification so far has relied on the second-order statistics of the available data $\mathbf{x}(n)$. Relying on the different frequency differences, the channel coefficients were identified by subspace methods. Next, we propose two methods that yield near-perfect channel estimates at high SNR's, require very few data points to be operational, and allow for coded inputs or inputs with unknown color.

IV. Deterministic Methods

In this section, we will adopt the channel model in (10) with $\omega_q := 2\pi q/N$, i.e., a set of harmonic frequencies. This introduces orthogonality to the complex exponential bases, which will be exploited in the two block-OFDM systems that will be proposed in this section to convert the matrix time-selective model in (5) to a multivariate convolutional relationship. In a different context, relying on the sampling theorem, harmonic FFT frequencies were also used in [19] for modeling channel time variation in code division multiple access (CDMA) systems. One should also keep in mind that any set of Q frequencies can be approximated arbitrarily well as $\omega_q = 2\pi q/N$ with a sufficiently large choice of N and appropriate choice of the index q.

A. Method I

Let the information symbol sequence be denoted by $s(n)$, and recall that $\mathbf{s}(n) := [s(nM)\cdots s(nM+M-1)]^T$. We conceive length $N-Q$ blocks of the $M \times 1$ vector sequence $\mathbf{s}(n)$ by defining $\mathbf{s}_p(n) := \mathbf{F}\mathbf{s}(p(N-Q)+n), n = 0, \cdots, N-Q-1$. Define the N-point IFFT of $\mathbf{s}_p(n)$

$$\overline{\mathbf{s}}_p(n) := \sum_{i=0}^{N-Q-1} \mathbf{s}_p(i)e^{j2\pi in/N}, \qquad n = 0, \cdots N-1$$

and use it as the blocked $M \times 1$ input to the filterbank precoder in Fig. 2. Therefore, for the pth block, we effectively transform the $M \times (N-Q)$ matrix $[\mathbf{s}_p(0)\cdots\mathbf{s}_p(N-Q-1)]$ into an $M \times N$ matrix $[\overline{\mathbf{s}}_p(0)\cdots\overline{\mathbf{s}}_p(N-1)]$.

In the noise-free case, the I/O relationship becomes $\mathbf{x}(n) = \sum_{q=0}^{Q} \boldsymbol{\Delta}_q\mathbf{H}_q\overline{\mathbf{s}}_p(n)e^{j2\pi qn/N}$. The channel identification will be performed block by block, where p is the block index for the vector sequence $\overline{\mathbf{s}}_p(n)$.

We know from (5) that the output vector sequence will be given by $\mathbf{x}_p(n) = \mathbf{H}(n)\overline{\mathbf{s}}_p(n)$. Taking the N-point componentwise FFT of both sides, we get

$$\begin{aligned} \mathbf{X}_p(l) &= \frac{1}{N}\sum_{n=0}^{N-1} \mathbf{H}(n)\overline{\mathbf{s}}_p(n)e^{-j2\pi nl/N} \\ &= \sum_{i=0}^{N-Q-1}\sum_{q=0}^{Q} \overline{\mathbf{H}}_q\mathbf{s}_p(i) \\ &\quad \cdot \left(\frac{1}{N}\sum_{n=0}^{N-1} \exp\left[j\frac{2\pi n}{N}(i+q-l)\right]\right) \end{aligned} \tag{18}$$

where, as in the previous sections, $\overline{\mathbf{H}}_q := \boldsymbol{\Delta}_q\mathbf{H}_q$, and $\boldsymbol{\Delta}_q := \text{diag}(1, e^{j2\pi q/NP}, \cdots, e^{j2\pi q(P-1)/NP})$. Substituting

$$\begin{aligned} &(1/N)\sum_{n=0}^{N-1} \exp\left[j\frac{2\pi n}{N}(i+q-l)\right] \\ &\quad = \sum_k \delta(i+q-l-kN) \end{aligned}$$

in (18), due to the limits of the sums, only the term corresponding to $k = 0$ survives. Thus, (18) becomes

$$\mathbf{X}_p(l) = \sum_{q=0}^{Q} \overline{\mathbf{H}}_q\mathbf{s}_p(l-q), \qquad l = 0, \cdots, N-1 \tag{19}$$

and $\mathbf{s}_p(i)=0$ for $i<0$ and $i>N-Q-1$. This MIMO I/O relation can be written in matrix form as

$$\underbrace{\begin{bmatrix} \mathbf{X}_p(N-1) & \cdots & \mathbf{X}_p(K) \\ \vdots & \cdots & \vdots \\ \mathbf{X}_p(N-K-1) & \cdots & \mathbf{X}_p(0) \end{bmatrix}}_{:=\mathcal{X}_p} = \underbrace{\begin{bmatrix} \overline{\mathbf{H}}_0 & \cdots & \overline{\mathbf{H}}_Q & \cdots & \mathbf{0} \\ \mathbf{0} & \ddots & \ddots & \ddots & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & \cdots & \overline{\mathbf{H}}_0 & \cdots & \overline{\mathbf{H}}_Q \end{bmatrix}}_{:=\mathcal{H}_{P(K+1)\times M(Q+K+1)}} \times \underbrace{\begin{bmatrix} \mathbf{s}_p(N-1) & \cdots & \mathbf{s}_p(K) \\ \vdots & \cdots & \vdots \\ \mathbf{s}_p(N-Q-K-1) & \cdots & \mathbf{s}_p(-Q) \end{bmatrix}}_{\mathcal{S}_p} \tag{20}$$

where K is a smoothing factor chosen to make $\mathcal{H}$ in (20) a tall matrix. The structure of $\mathcal{H}$ in (20) suggests a subspace approach to estimate $\mathbf{H}_q$ deterministically.

We require the following assumptions for the identifiability of $\{\overline{\mathbf{H}}_q\}_{q=0}^{Q}$ from $\{\mathbf{X}_p(l)\}_{l=0}^{N-1}$.

- a3) Matrix $\mathcal{H}$ in (20) is full column rank. For this condition to hold, it is necessary that $\{H_l(z\rho^l)\}_{l=0}^{L}$ are coprime, where $H_l(z):=\sum_{q=0}^{Q}h_q(l)z^{-q}$, and $\rho:=\exp(-2\pi/NP)$ (see Appendix B for the proof). This necessary condition is a useful characterization of whether $\mathcal{H}$ is ill conditioned in terms of the channel coefficients.
- a4) The input is persistently exciting enough to assure that $\mathcal{S}_p$ in (20) is full rank.
- a5) The number of blocks is selected to satisfy $N\geq Q+2K+1$ (and, thus, $\mathcal{S}_p$ is fat).

Recall that K is chosen to make $\mathcal{H}$ tall. This means that the bigger P/M is (i.e., the more redundancy we allow), the smaller K needs to be to satisfy $P(K+1)\geq M(Q+K+1)$ (the tallness condition on $\mathcal{H}$). Therefore, if the redundancy is big, K can be chosen small, and Q can afford to be closer to N, which shows that with higher redundancy, the method can tolerate channels with faster variation [a larger Q in (7)].

Let the SVD of $\mathcal{X}_p\mathcal{X}_p^H$ be given by

$$\mathcal{X}_p\mathcal{X}_p^H=[\overline{\mathbf{U}}\quad\tilde{\mathbf{U}}]\begin{bmatrix}\mathbf{\Sigma}&\mathbf{0}\\ \mathbf{0}&\mathbf{0}\end{bmatrix}\begin{bmatrix}\overline{\mathbf{U}}^H\\ \tilde{\mathbf{U}}^H\end{bmatrix}. \tag{21}$$

By a3)–a5), we can write $\mathcal{R}(\mathcal{X}_p\mathcal{X}_p^H)=\mathcal{R}(\mathcal{H})$. Since $\tilde{\mathbf{U}}^H\mathcal{X}_p\mathcal{X}_p^H=\mathbf{0}$, it follows that $\tilde{\mathbf{u}}_l^H\mathcal{H}=\mathbf{0}^H$, where $\tilde{\mathbf{u}}_l:=[\tilde{\mathbf{u}}_l^T(0)\cdots\tilde{\mathbf{u}}_l^T(K)]^T$ is the lth column of $\tilde{\mathbf{U}}$. Since $\mathcal{H}$ is a block convolution matrix, the orthogonality condition $\tilde{\mathbf{u}}_l^H\mathcal{H}=\mathbf{0}^H$ can be written as

$$[\overline{\mathbf{H}}_0^H\cdots\overline{\mathbf{H}}_Q^H]\tilde{\mathcal{U}}=\mathbf{0} \tag{22}$$

where $\tilde{\mathcal{U}}:=[\tilde{\mathbf{U}}_1\cdots\tilde{\mathbf{U}}_{P(K+1)-M(Q+K+1)}]$, and $\tilde{\mathbf{U}}_l$ is a $P(Q+1)\times(Q+K+1)$ matrix given by

$$\tilde{\mathbf{U}}_l:=\begin{bmatrix}\tilde{\mathbf{u}}_l(0)&\cdots&\tilde{\mathbf{u}}_l(K)&\cdots&\mathbf{0}\\ \mathbf{0}&\cdots&\vdots&\ddots&\mathbf{0}\\ \vdots&\ddots&\vdots&\vdots&\vdots\\ \mathbf{0}&\cdots&\tilde{\mathbf{u}}_l(K-Q)&\cdots&\tilde{\mathbf{u}}_l(K)\end{bmatrix}. \tag{23}$$

It can be shown [1] that by using (22), one can obtain $\mathbf{T}[\overline{\mathbf{H}}_0^H\cdots\overline{\mathbf{H}}_Q^H]$ from the output $\mathbf{X}_p(l)$ in (19), where $\mathbf{T}$ is an $M\times M$ matrix ambiguity. This means that using output only data, for every q, we can estimate $\mathbf{T}\mathbf{H}_q^H\mathbf{\Delta}_q^H$, which, upon multiplying with $\mathbf{\Delta}_q^{-H}$ from the right, gives us a basis for the row span of $\mathbf{H}_q^H$. It was shown in Section III that the row span of $\mathbf{H}_q^H$ provides enough information to estimate the coefficients inside it.

As was the case for the statistical method, estimating each $\mathbf{H}_q$ separately will create a scalar ambiguity β_q for each q. However, notice that $\mathbf{T}[\mathbf{H}_0^H\cdots\mathbf{H}_Q^H]$, which we can obtain via $\mathbf{T}[\overline{\mathbf{H}}_0^H\cdots\overline{\mathbf{H}}_Q^H]\,\mathrm{diag}(\mathbf{\Delta}_0^{-H}\cdots\mathbf{\Delta}_Q^{-H})$, can be estimated from the knowledge of its row span because $[\mathbf{H}_0^H\cdots\mathbf{H}_Q^H]^H$ has the same banded Toeplitz structure as each $\mathbf{H}_q$.

In a nutshell, the steps involved in estimating $\mathbf{H}_q$ deterministically are as follows.

- Step 1) Transform the $M\times 1$ vectorized *and* blocked sequence $\mathbf{s}_p(n):=\mathbf{F}\mathbf{s}(p(N-Q)+n)$ by taking a componentwise IFFT.
- Step 2) Compute $\mathbf{X}_p(l)$ using (18).
- Step 3) Form $\mathcal{X}_p$ in (20), and estimate $\tilde{\mathbf{U}}$ from the SVD of $\mathcal{X}_p\mathcal{X}_p^H$.
- Step 4) Using (22) and (23), form $\tilde{\mathcal{U}}$, and estimate the row space of $[\overline{\mathbf{H}}_0^H\cdots\overline{\mathbf{H}}_Q^H]$ from (22).
- Step 5) Using the structure of $\mathbf{H}_q$, estimate it from its column span, as described in Section III.

The input symbols in $\mathcal{S}_p$ can be estimated as $\hat{\mathcal{S}}_p=\hat{\mathcal{H}}^{\dagger}\mathcal{X}_p$. The bandwidth efficiency, which is defined as the ratio of the number of symbols in the input and output blocks, is given by

$$\mathcal{E}_I=\frac{M(N-Q)}{(M+L)N}.$$

For small values of Q and L, this ratio is close to one. Next, we propose a different precoding scheme that performs better and is insensitive to channel root locations at the expense of a lower bandwidth efficiency.

B. Method II

Due to its structure, $\mathcal{H}$ in the previous subsection is susceptible to losing rank. Even though, as argued for the output diversity methods, it may be unlikely to have a combination of channel coefficients for $\mathcal{H}$ to lose rank, Method I of the previous subsection suffers from $\mathcal{H}$ being ill conditioned, as is also verified in our simulations, particularly due to noise amplification. Here, we will generalize the approach in [21] to jointly time- and frequency-selective channels and show that the resulting channel matrix does not suffer from a disparity condition, i.e., identifiability is guaranteed irrespective of the subchannel zero locations.

As in the previous subsection, suppose that the channel variation is modeled by harmonic frequencies so that $\omega_q := 2\pi q/(N_2P_2)$, where $N_2P_2 = N$. Let $\mathbf{s}_p(n) := \mathbf{F}\mathbf{s}(pN_2M_2+n)$ denote a length-N_2M_2 block of $M\times 1$ vectors. Generalizing the OFDM concept in another direction, we perform the following transformation on the blocked input sequence $\mathbf{s}_p(n)$ to obtain the precoded sequence $\overline{\mathbf{s}}_p(n)$

$$\overline{\mathbf{s}}_p(n) := \sum_{i=0}^{N_2-1}\sum_{m=0}^{M_2-1} \mathbf{s}_p(iM_2+m) \cdot \exp[j2\pi(iP_2+m)n/(N_2P_2)] \tag{24}$$

where $P_2 = M_2+Q$, and $M_2 > Q$. Notice that the transformation in (24) is the row-by-row IFFT of the matrix $[\mathbf{s}_p(0)\cdots\mathbf{s}_p(M_2-1)\ \mathbf{0}_{M\times Q}\mathbf{s}_p(P_2)\cdots\mathbf{s}_p(P_2+M_2)\ \mathbf{0}_{M\times Q}\mathbf{s}_p((N_2-1)P_2+M_2)\ \mathbf{0}_{M\times Q}]_{M\times N_2P_2}$.

Suppose we take the N_2P_2-point FFT of the received sequence $\mathbf{H}(n)\overline{\mathbf{s}}_p(n)$:

$$\mathbf{X}_p(l) = \frac{1}{N_2P_2}\sum_{n=0}^{N_2P_2-1} \mathbf{H}(n)\overline{\mathbf{s}}_p(n)e^{-j2\pi nl/N_2P_2}. \tag{25}$$

Substituting (24) into (25), we obtain

$$\mathbf{X}_p(l) = \sum_{i=0}^{N_2-1}\sum_{m=0}^{M_2-1}\sum_{q=0}^{Q} \overline{\mathbf{H}}_q\mathbf{s}_p(iM_2+m)\,\frac{1}{N_2P_2} \cdot\left(\sum_{n=0}^{N_2P_2-1}\exp[j2\pi(iP_2+m+q-l)n/N_2P_2]\right) \tag{26}$$

but we know that the term in parentheses equals $\sum_j \delta(iP_2+m+q-l-jN_2P_2)$. Given that $i=0,\cdots,N_2-1$, $m=0,\cdots,M_2-1$, $q=0,\cdots,Q$, and letting $l=kP_2+v$, the triple sum in (26) can be reduced to a double sum by setting $j=0$ and $i=k$, which yields

$$\begin{aligned}\mathbf{X}_p(kP_2+v) &= \sum_{m=0}^{M_2-1}\sum_{q=0}^{Q}\overline{\mathbf{H}}_q\mathbf{s}_p(kM_2+m)\delta(m+q-v)\\ &= \sum_{m=0}^{M_2-1}\overline{\mathbf{H}}_{v-m}\mathbf{s}_p(kM_2+m).\end{aligned} \tag{27}$$

Writing (27) in matrix form, we obtain

$$\underbrace{\mathcal{X}_p}_{PP_2\times N_2} = \underbrace{\mathcal{H}}_{PP_2\times MM_2}\ \mathcal{S}_p \tag{28}$$

where $\mathcal{X}_p((j-1)P+1:jP,i)=\mathbf{X}_p(iP_2+j)$, $\mathcal{S}_p((j-1)M+1:jM,i)=\mathbf{s}_p(iM_2+j)$, and

$$\mathcal{H} := \begin{bmatrix} \overline{\mathbf{H}}_0 & \mathbf{0} & \cdots & \mathbf{0}\\ \vdots & \overline{\mathbf{H}}_0 & \ddots & \vdots\\ \overline{\mathbf{H}}_Q & \cdots & \ddots & \mathbf{0}\\ \mathbf{0} & \ddots & \cdots & \overline{\mathbf{H}}_0\\ \vdots & \ddots & \ddots & \vdots\\ \mathbf{0} & \cdots & \mathbf{0} & \overline{\mathbf{H}}_Q \end{bmatrix},\quad \overline{\mathbf{H}}_q := \boldsymbol{\Delta}_q\mathbf{H}_q \tag{29}$$

and $\mathbf{H}_q$ is given by (11).

Due to its upper triangular structure, $\mathcal{H}$ has full rank. Therefore, provided that

a6) $\mathcal{S}_p$ is full rank,

a7) $\mathcal{S}_p$ is fat (satisfied by choosing $N_2 \geq MM_2$),

we have $\mathcal{R}(\mathcal{X}_p\mathcal{X}_p^H)=\mathcal{R}(\mathcal{H})$. From the singular value decomposition of $\mathcal{X}_p\mathcal{X}_p^H$, we can obtain a basis for the orthogonal complement of $\mathcal{R}(\mathcal{H})$, which is given by $\tilde{\mathbf{U}}$. Let $\tilde{\mathbf{U}}(i)$ be matrices of the same size satisfying $\tilde{\mathbf{U}}^H=[\tilde{\mathbf{U}}^H(0)\cdots\tilde{\mathbf{U}}^H(P_2-1)]$. From the orthogonality condition $\tilde{\mathbf{U}}^H\mathcal{H}=\mathbf{0}$, generalizing the result in Section III, we obtain $\mathbf{T}[\overline{\mathbf{H}}_0^H\cdots\overline{\mathbf{H}}_Q^H]$, where $\mathbf{T}$ is an $M\times M$ matrix ambiguity, by solving the equation $\mathbf{H}_T\mathcal{U}=\mathbf{0}$ for $\mathbf{H}_T$, where $\mathcal{U}$ is a block Hankel matrix whose first block column is $[\tilde{\mathbf{U}}^T(0)\cdots\tilde{\mathbf{U}}^T(Q)]^T$, and whose last block row is $[\tilde{\mathbf{U}}(Q)\cdots\tilde{\mathbf{U}}(P_2-1)]$. Since $\boldsymbol{\Delta}_q$'s are known, we can obtain, upon multiplication, $\mathbf{T}[\mathbf{H}_0^H\cdots\mathbf{H}_Q^H]$. However, we know from Section III that the ambiguity matrix $\mathbf{T}$ can be resolved due to the structure of $\mathbf{H}_q$.

Given Q, choose $M_2 > Q$ and $P_2 = M_2+Q$. Then, choose N_2 such that the frequencies $\overline{\omega}_q$ will fall sufficiently closely to the grid $2\pi q/(N_2P_2P)$, where $P=M+L$, and $M>L$. Hence, the choice of N_2 provides a tradeoff between computational complexity (because it affects the sizes of the matrices involved in the estimation) and model accuracy because the larger N_2 is, the finer the grid $2\pi q/(N_2P_2P)$ will be. Given these system parameters, here are the steps involved for the second deterministic method.

Step 1) Use (24) to obtain $\overline{\mathbf{s}}_p(n)$.

Step 2) Calculate $\mathbf{X}_p(l)$ using (25).

Step 3) Construct $\mathcal{X}_p$, and using its SVD, construct $\mathcal{U}$.

Step 4) From the left nullspace of $\mathcal{U}$, obtain $\mathbf{T}[\mathbf{H}_0^H\boldsymbol{\Delta}_0^H\cdots\mathbf{H}_Q^H\boldsymbol{\Delta}_Q^H]$, and knowing $\boldsymbol{\Delta}_q$, calculate $\mathbf{T}[\mathbf{H}_0^H\cdots\mathbf{H}_Q^H]$.

Step 5) Using the structure of $\mathbf{H}_q$, resolve the matrix ambiguity $\mathbf{T}$, as was done in Section I. The problem of resolving the Q scalar ambiguities β_q corresponding to each $\mathbf{H}_q$ can be resolved by estimating $[\mathbf{H}_0^H\cdots\mathbf{H}_Q^H]^H$ jointly, as mentioned in the previous section.

Similar to the previous subsection, the input symbols can be estimated as $\hat{\mathcal{S}}_p=\hat{\mathcal{H}}^{\dagger}\mathcal{X}_p$. Notice that there is no channel disparity condition here since $\mathcal{H}$ is always full rank. The price paid is the decrease in the bandwidth efficiency, which is given by

$$\mathcal{E}_{II} = \frac{MM_2}{PP_2} = \frac{MM_2}{(M+L)(M_2+Q)}.$$

Remark 3: Both Methods I and II recover time *and* frequency selective channels, whereas the methods in [21] and [12] are blind methods for flat fading (time-selective) channels only.

V. Symbol Recovery

In this section, we summarize several symbol recovery techniques that could be used after the channel has been estimated, particularly for the statistical method. After describing a method that relies on parameterizing the null space of a channel matrix, we will go over the zero-forcing (ZF) and minimum mean-

square error (MMSE) alternatives. At the end of the section, we will address the problem of symbol recovery for the deterministic methods.

Having estimated $\{\mathbf{H}_q\}_{q=0}^{Q}$ and having designed $\mathbf{F}$, we let $\mathcal{H}_F := [\boldsymbol{\Delta}_0\mathbf{H}_0\mathbf{F}\cdots\boldsymbol{\Delta}_Q\mathbf{H}_Q\mathbf{F}]$, $\mathbf{s}(n)$ be the $M\times 1$ input block, and $\mathbf{s}_b(n) := [e^{j\omega_0 n}\mathbf{s}^T(n)\cdots e^{j\omega_Q n}\mathbf{s}^T(n)]^T$. Then, (10) can be written as follows:

$$\mathbf{x}(n) = \underbrace{\mathcal{H}_F}_{P\times M(Q+1)}\mathbf{s}_b(n) + \mathbf{v}(n). \tag{30}$$

We wish to solve (30) for $\mathbf{s}_b(n)$, having knowledge of the output $\mathbf{x}(n)$ and $\mathcal{H}_F$. Because $\mathcal{H}_F$ is a fat matrix, the solution is not unique unless we use the structure of $\mathbf{s}_b(n)$. The solutions of (30) for $\mathbf{s}_b(n)$ can be parameterized by the product $\mathbf{s}_b(n) := \mathbf{V}(n)\boldsymbol{\theta}(n)$, where the first column of the $(Q+1)M\times(Q+1)M-P+1$ matrix $\mathbf{V}(n)$ is the particular solution $\mathcal{H}_F^{\dagger}\mathbf{x}(n)$, and the other columns span $\mathcal{N}(\mathcal{H}_F)$ (hereafter the dependence of $\boldsymbol{\theta}$ and $\mathbf{V}$ on n will be dropped for brevity). The task here is to identify the parameter $\boldsymbol{\theta}$ so that $\mathbf{V}\boldsymbol{\theta} = [e^{j\omega_0 n}\cdots e^{j\omega_Q n}]^T\otimes\hat{\mathbf{s}}(n)$. Let $\mathbf{V}_q$ be $M\times(Q+1)M-P+1$ submatrices of $\mathbf{V} = [\mathbf{V}_0^T\cdots\mathbf{V}_Q^T]^T$. Then, solving for $\boldsymbol{\theta}$ under the structural constraint on $\mathbf{s}_b(n)$ is equivalent to solving the system $\mathbf{V}_q e^{-j\omega_q n}\boldsymbol{\theta} = \mathbf{V}_{q+1}e^{-j\omega_{q+1}n}\boldsymbol{\theta}$, $q = 0,\cdots,Q-1$. After solving for $\boldsymbol{\theta}$, we can compute $\mathbf{V}\boldsymbol{\theta}$ to obtain $\mathbf{s}_b(n)$.

An alternative Viterbi-like approach would be to use the finite alphabet property of $s(n)$ for every n to construct all possible $\mathbf{s}_b(n)$'s and choose the one that comes closest to satisfying (30) in the least squares sense. The computational complexity of this procedure, however, is $\mathcal{O}(|\mathcal{C}|^M)$, where $|\mathcal{C}|$ is the size of the signal constellation.

Computationally simpler input recovery can also be established through zero-forcing or MMSE linear equalizers. For this purpose, consider the I/O relation in (5). The zero-forcing solution for this scheme is given by $\hat{\mathbf{s}}(n) = (\mathbf{H}(n)\mathbf{F})^{\dagger}\mathbf{x}(n)$. It is well known that zero-forcing schemes do not take the noise into account and suffer at moderate or low SNR's. This motivates us to consider an equalizer matrix $\mathbf{G}(n)$ that minimizes the following MSE:

$$\hat{\mathbf{G}}(n) = \arg\min_{\mathbf{G}(n)} E|\mathbf{G}(n)\mathbf{x}(n) - \mathbf{s}(n)|^2. \tag{31}$$

Using the orthogonality principle, it follows readily that the solution to this problem is given by

$$\hat{\mathbf{G}}(n) = \mathbf{R}_{ss}\mathbf{F}^H\mathbf{H}^H(n)(\mathbf{R}_{vv} + \mathbf{H}(n)\mathbf{F}\mathbf{R}_{ss}\mathbf{F}^H\mathbf{H}^H(n))^{-1}$$

where $\mathbf{R}_{vv} := E[\mathbf{v}(n)\mathbf{v}^H(n)]$. The common problem in all input-recovery schemes for the statistical method is that some sort of inversion has to take place for each block index n, which is computationally expensive. This problem could be alleviated by approximating the frequencies ω_q on a FFT grid of size N, which amounts to assuming that $\mathbf{H}(n)$ is (approximately) periodic with period N, hence necessitating the inversion of $\mathbf{H}(n)$ only N times.

By approximating the frequencies of the BEM on the FFT grid, in the deterministic Methods I and II, we convert a SISO model into a MIMO one. The equalization problem in this case is equivalent to the inversion of $\mathcal{H}$, which is independent of the time index n. This means that for the deterministic methods, we need to invert the channel matrix $\mathcal{H}$ only when the channel parameters have changed significantly.

In this paper, we have assumed $\mathbf{F}$ to be an arbitrary full-rank matrix. For the case where the channel is time invariant, precoder design issues (choice of $\mathbf{F}$) have been dealt with in [20]. How the choice of $\mathbf{F}$ influences the bit error rate is discussed in the simulations.

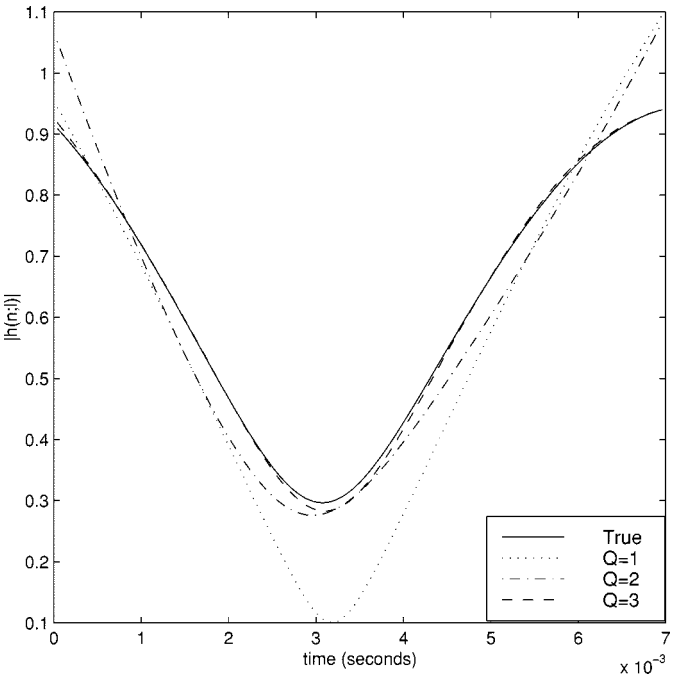

Fig. 4. Model validation.

VI. Simulations

In this section, we corroborate the ideas discussed in this paper with computer simulations. We will first provide simulations illustrating the validity and limitations of the model and then simulations for the statistical method, and we will illustrate and compare the results for the deterministic methods. Finally, we will depict the performance of one of the deterministic methods in the presence of model mismatch. In all simulations, the input constellation used was equiprobable quadrature amplitude modulation (4-QAM) with unit variance.

A. Basis Expansion Model

In Fig. 4, we examine how accurately the complex exponential basis can capture the channel variation for realistic block lengths, bit rates, mobile speeds, and carrier frequencies for a single channel tap whose variation is characterized by the Jakes' spectrum. For a system with a carrier frequency of 900 MHz, bit rate of 24 KHz, and block length of 170 symbols, we have generated a channel tap $|h(n;l)|$ shown with the solid line, using the formula

$$h(n;l) = c(l)\frac{1}{\sqrt{K}}\sum_{k=0}^{K-1}\cdot\exp[j(2\pi f_{D_{\max}}\cos(\alpha_{k,l})n + \phi_{k,l})].$$

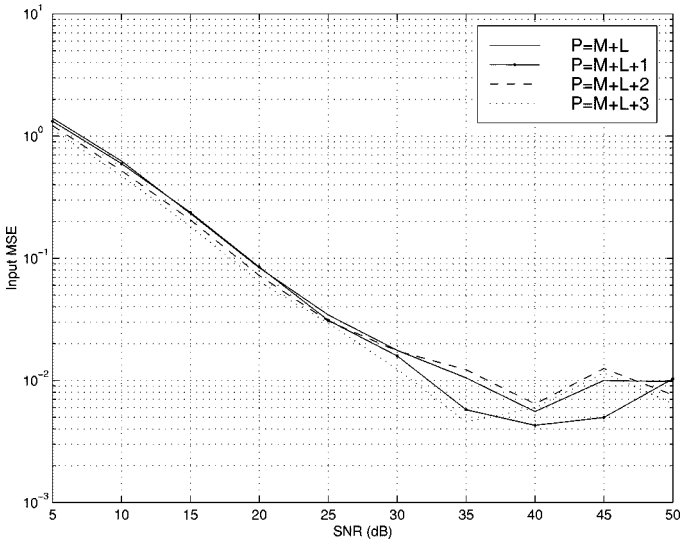

Fig. 5. Order overestimation.

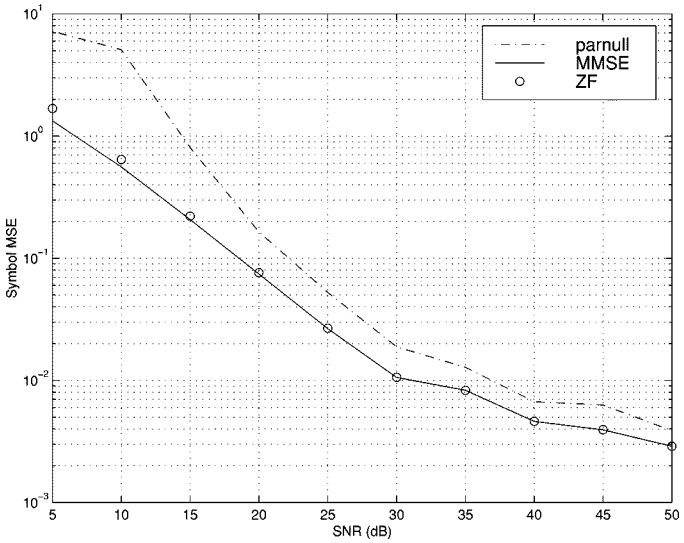

Fig. 6. Different equalization schemes.

where $c(l) = 1$ controls the power of the lth tap chosen according to the power delay profile, $f_{D_{\max}} = 66$ Hz (corresponding to a 80 km/hr speed) is the desired maximum Doppler frequency, $K = 100$ is the number of sinusoids, and $\alpha_{k,l}$, $\phi_{k,l}$ are mutually independent, uniformly distributed random variables. It is well known (see, e.g., [8]) that such a sum will yield a random process $h(n;l)$ in the time index n, whose power spectrum will approximate the Jakes' spectrum arbitrarily well with increasing K. We labeled this realization of $h(n;l)$ "True" and approximated $h(n;l)$ with a basis expansion $\hat{h}(n;l) = \sum_{q=0}^{Q} h_q(l) e^{j2\pi f_{D_{\max}} qn/Q}$ to minimize (1/170) $\sum_{n=0}^{169} |h(n;l) - \hat{h}(n;l)|^2$ with respect to $h_q(l)$ for $Q = 1, 2, 3$. We observe that as Q increases, the approximation improves, and for $Q = 3$, it is almost perfect for this block length ($N = 170$), which is a typical block length in practice. Similar results were obtained for all realizations of $h(n;l)$, even when the frequencies were not chosen to be equispaced as above ($\omega_q \neq 2\pi f_{D_{\max}} q/Q$). Due to the narrowband nature of the variation, complex exponentials prove to be good basis functions for practical block lengths.

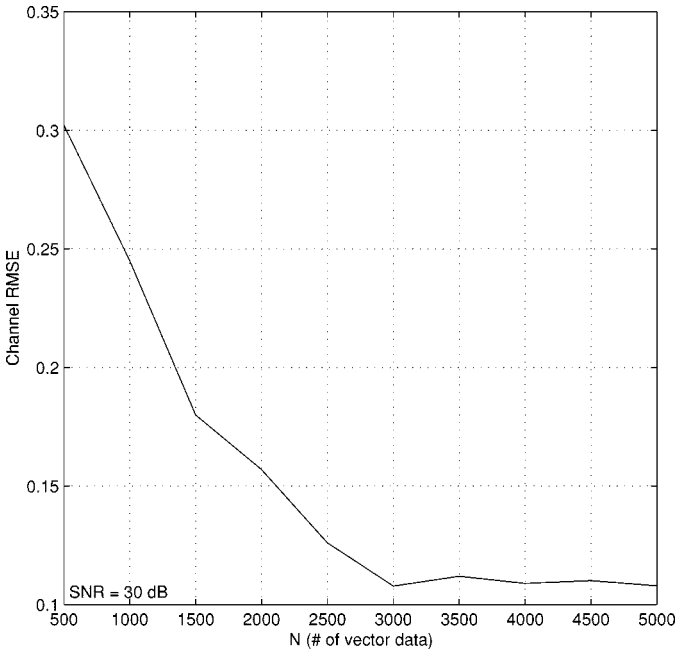

Fig. 7. Improvement with data length.

B. Statistical Method

In Figs. 5–8, the parameters were chosen as $M = 5$ and $L = 3$, the precoder matrix was $\mathbf{F} = \mathbf{I}$, and the frequencies were $\overline{\omega}_0 = 0$ (modeling the TI part of the channel), $\overline{\omega}_1 = 2\pi/40$, and $\overline{\omega}_2 = 2\pi/104$. These frequencies correspond to a vehicle speed of 115 km/h for a system operating at 1.8 GHz and a bit rate of 20 KHz.

In Fig. 5, we illustrate how channel order overestimation does not change and actually slightly reduces the symbol mean square error (MSE) with usage of larger blocks of size $P = M + \overline{L} = M + L + i$, $i \in [0, 3]$, at low SNR's. In Fig. 5, $N = 5000$ data were used.

In Fig. 6, the three input recovery schemes are compared in terms of MSE between the true and the equalized inputs. It is observed that the MMSE equalizer is the best among the three proposed linear equalizers for all SNR's and that the ZF scheme as well as the method that parameterizes the nullspace (parnull) approximates the MMSE closely at high SNR's. In Fig. 6, $N = 5000$ data were used, which are not many, considering the explicit modeling of the time variation captured by the basis expansion.

Fig. 7 illustrates how at SNR $= 30$ dB, increasing the data length yields reduced channel root mean square error (RMSE), illustrating its consistency and showing the superiority of this averaging-based method over its deterministic counterparts. Fig. 7 was averaged over 200 realizations of the RMSE curve.

In Fig. 8, we used the MMSE equalizer with $N = 1000$ data points to obtain a symbol error rate (SER) plot. The plot was obtained by counting errors across 500 000 equalized symbols for each value of SNR. The simulations suggest that the statistical method is to be preferred when the SNR is relatively low, and a relatively long data record (during which the channel param-

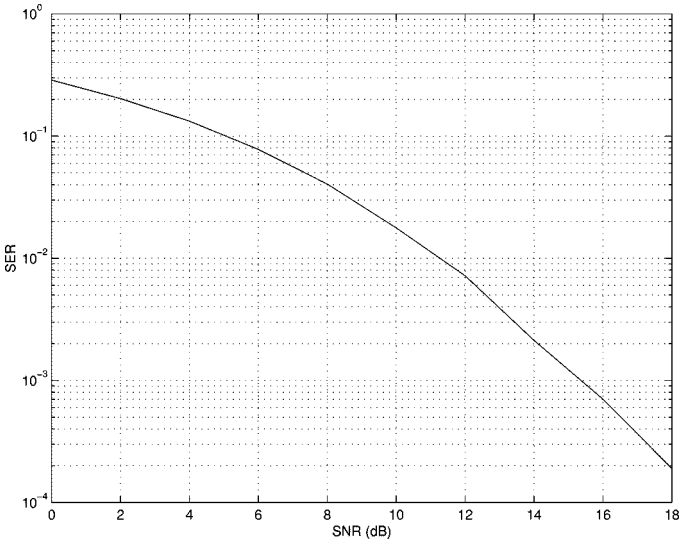

Fig. 8. SER for the MMSE equalizer (stat. meth.)

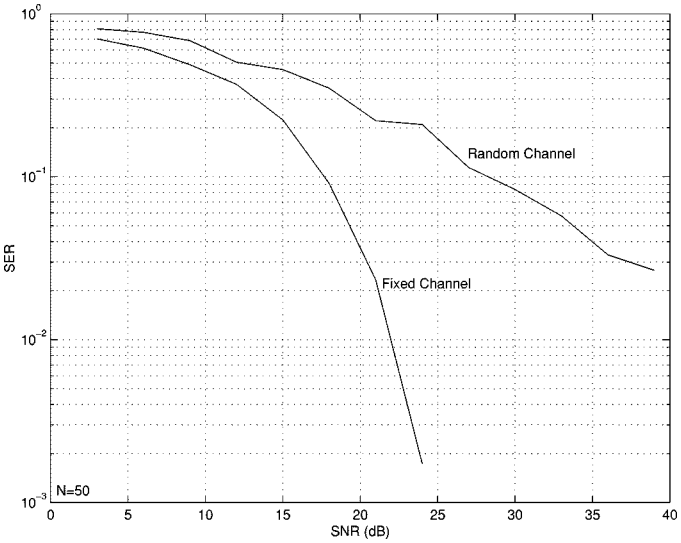

Fig. 10. SER versus SNR (det. meth. I).

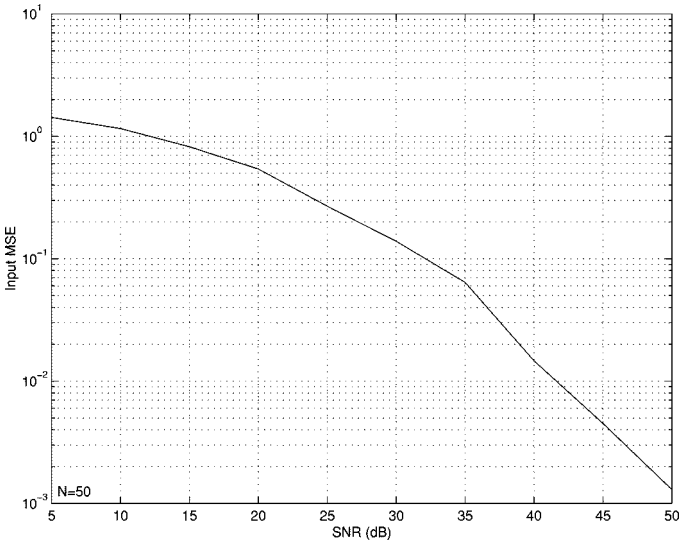

Fig. 9. Input MSE versus SNR (det. meth. I).

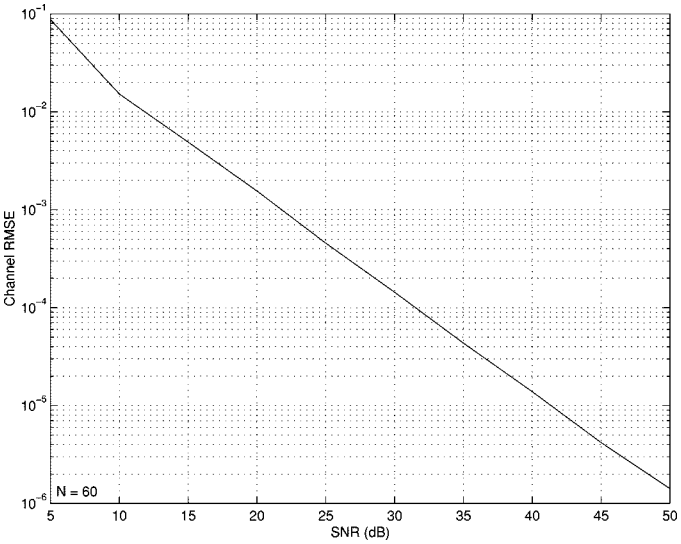

Fig. 11. Channel RMSE versus SNR (det. meth. II).

eters have not changed) is available. This assumption is reasonable since the channel variation is explicitly modeled through the BEM. In the next subsection, we illustrate the deterministic methods that require an order of magnitude less data at high SNR's.

C. Deterministic Methods

To illustrate the first deterministic method, we used the channel estimation algorithm with (Q, M, L, P, K) = (2, 4, 3, 7, 5). The frequencies were $\overline{\omega}_0 = 0$, $\overline{\omega}_1 = 2\pi/350$, and $\overline{\omega}_2 = 4\pi/350$, which corresponds to a vehicle speed of about 70 km/h at a carrier frequency of 1.8 GHz and a bit rate of 20 KHz.

It should be noted that for the first deterministic method, different realizations of the channel gave different results for the same SNR due to fact that the zero locations of the channel influence the condition number of the channel matrix $\mathcal{H}$ and, hence, cause noise amplification through $\mathcal{H}^{\dagger}$.

Fig. 9 shows the MSE between the true and estimated inputs versus SNR with $N = 50$ snapshots available. Fig. 9 illustrates that the increase in the SNR for the deterministic Method I makes a more significant improvement than it does for the statistical method because here, noise is the only factor that causes estimation errors.

In Fig. 10, we computed the symbol error rate to shed light on the performance of the deterministic Method I. The curve marked "random channel" is obtained by performing an average over 100 randomly generated (Rayleigh faded) channels that are obtained when the coefficients were Gaussian with zero-mean and unit variance. The average performance is shown to be inferior to the performance where the channel coefficients are fixed, which illustrates the fact that Method I suffers from the channel disparity condition discussed in Appendix II. In this sense, Method I is similar to the second-order output diversity methods (e.g., [14]).

In Figs. 11 and 12, we test the performance of the deterministic Method II. In both Figs. 11 and 12, the parameters were

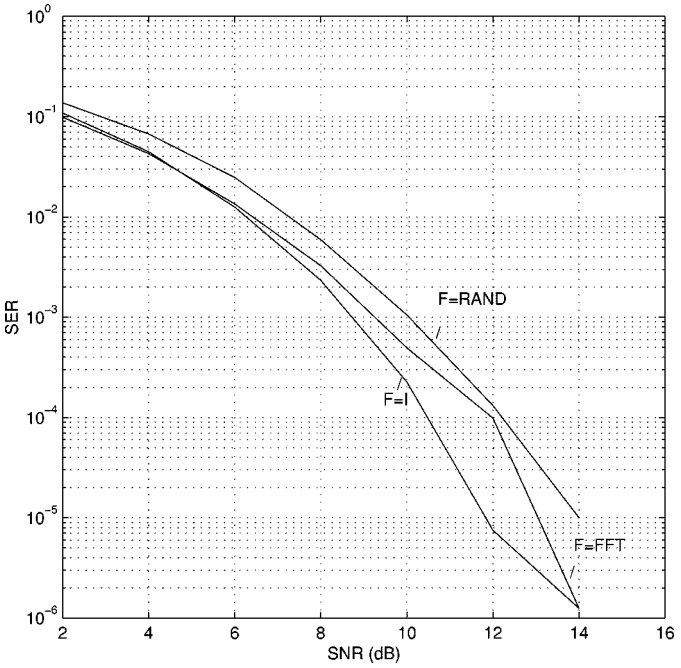

Fig. 12. SER versus SNR, colored inp. (det. meth. II).

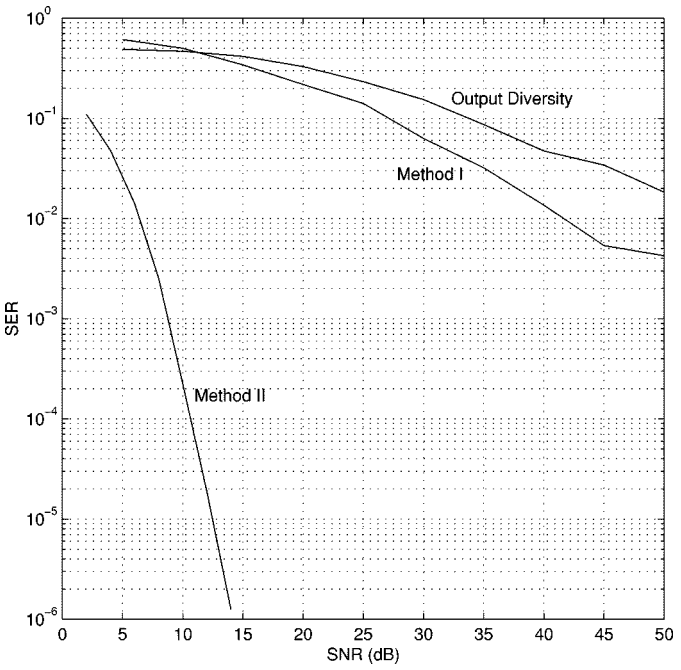

Fig. 13. SER versus SNR (comparisons).

$(Q, M, M_2, L, P, P_2) = (2, 4, 4, 3, 7, 6)$. Fig. 11 shows that with $N = 60$ snapshots, for high SNR's, the improvement of the channel RMSE is approximately exponential. In Fig. 12, we illustrate how the symbol error rate improves with SNR for colored and white inputs corresponding to different precoding matrices $\mathbf{F}$. The choice $\mathbf{F} = \mathbf{I}$ corresponds to a white input sequence, which has undergone no precoding, and $\mathbf{F} =$ FFT corresponds to an FFT matrix, which transforms the input but also does not introduce any color since it is unitary. Both matrices have a condition number of 1. Finally, we tried a randomly

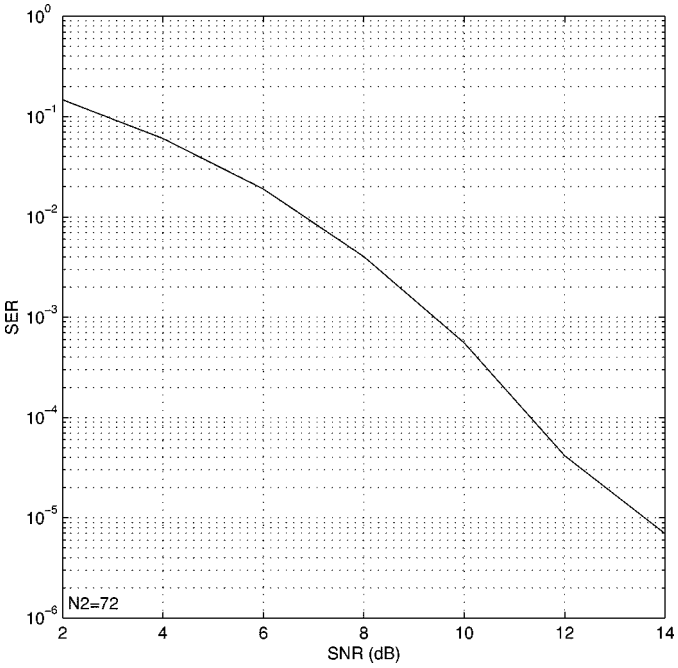

Fig. 14. SER versus SNR. (det. meth. II).

generated matrix that is denoted by $\mathbf{F} =$ RAND with a condition number close to 2. It was observed that a higher condition number of $\mathbf{F}$ caused a higher condition number of the realizations of $\mathcal{S}_p$, which increased the bit error rate. This is because with $\mathbf{F}$ close to losing rank, so is $\mathcal{S}_p$, and this violates a6). We conclude that if $s(n)$ is already white, $\mathbf{F} = \mathbf{I}$ is the recommended choice, whereas if $s(n)$ is colored, $\mathbf{F}$ could be chosen to decorrelate the input, which for most realizations of $\mathcal{S}_p$ will result in a low condition number.

Fig. 13 provides a comparison with the direct equalization method of [5], which exploits output (antenna) diversity. We computed the error probability over 100 000 equalized symbols for each SNR, where we averaged over different realizations of the channel taps to see, on the average, which methods suffer from a channel disparity condition. We observe that similar to the output diversity method, Method I suffers from the ill conditioning of the channel matrix but performs slightly better than the output diversity method. Performance of Method II, on the other hand, as was mentioned earlier, is not seen to be affected from the different realizations of the channel coefficients.

D. Performance with Mismatched BEM

In our last set of simulations, we calculated the performance of Method II for a channel whose time varying coefficients were generated according to the following sum:

$$h(n;l) = c(l)\,\frac{1}{\sqrt{K}}\sum_{k=0}^{K-1} \cdot \exp[j(2\pi f_{D_k} n + \phi_{k,l})]$$

where $\phi_{k,l}$ are mutually independent, uniformly distributed random variables, and f_{D_k} are the Doppler frequencies, the distribution of which determine the spectral characteristics of the channel taps. The maximum Doppler frequency

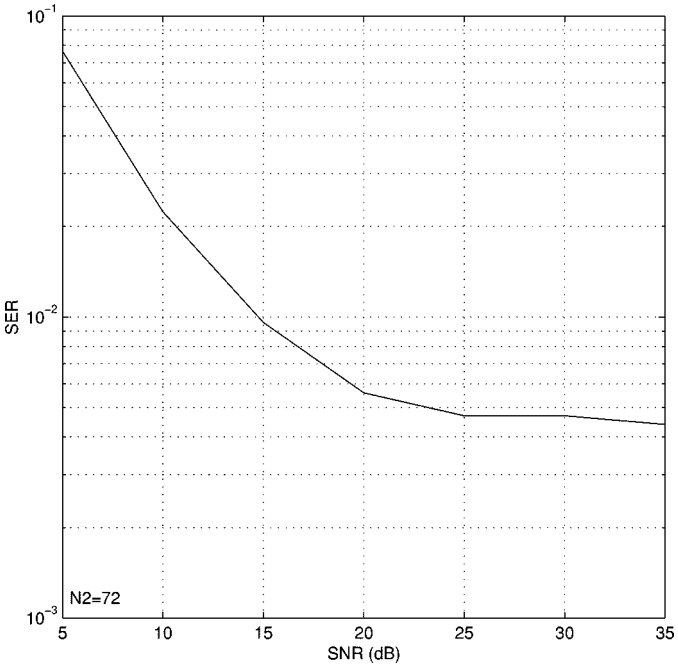

Fig. 15. SER versus SNR, imperfect model (meth. II).

$f_{D_{\max}} = \max_k f_{D_k}$ in Figs. 14 and 15 corresponds to a vehicle speed of $v = 50$ km/h, a bit rate of 24 KHz, and a carrier frequency of 900 MHz.

In Fig. 14, we plot the symbol error rate for a two-tap channel we generated with f_{D_k}, $k = 0, \cdots, 100$ chosen uniformly from the set $\{f_{D_{\max}} q/Q\}_{q=0}^{Q}$, which corresponds to the basis expansion model with harmonically spaced frequencies. In Fig. 15, for both taps, we generated the first 80 frequencies from the set $\{f_{D_{\max}} q/Q\}_{q=0}^{Q}$, and the last 20 were chosen from $f_{D_{\max}} \cos(\alpha_{k,l})$, which yields Jakes' spectrum superimposed with a spectrum that has harmonically related frequencies. Under these conditions, the SER improves until about 20 dB, and then exhibits an error floor, due to model mismatch created by the frequencies that are unaccounted for. Nevertheless, the method does not break down, despite the presence of these frequencies. The parameters for Figs. 14 and 15 were $(Q, M, M_2, L, P, P_2) = (3, 2, 5, 1, 3, 8)$ and $(Q, M, M_2, L, P, P_2) = (4, 2, 5, 1, 3, 8)$, respectively.

VII. Conclusions

We derived novel blind channel estimation and symbol recovery algorithms for time-varying communication channels where the variation is captured by a complex exponential basis expansion model. It was shown that for the statistical method, introducing redundancy at the input brings about many advantages that the output diversity methods do not possess, such as robustness to order overestimation, insensitivity to channel zero locations and stationary noise, and the fact that without oversampling, only one antenna output suffices for identifiability of time- and frequency-selective channels.

In low SNR settings where noise averaging is required, the statistical method in the first half of the paper should be used, whereas if the SNR is moderate/high, but quick channel acquisition with a few data points is required, then the methods in the second half of the paper should be favored.

The deterministic methods only require few data points for near-perfect estimation at high SNR's, do not suffer from estimation errors, and allow the input to be colored. We developed two methods where the tradeoff was shown to be between higher redundancy and better performance/less restrictive assumptions.

Appendix A

In this Appendix, we will show identifiability without assuming a2), removing any restrictive assumptions on the frequencies $\{\omega_q\}$. In general, without assuming a2), the cyclic correlation matrix is given by

$$\mathbf{C}_{xx}(\omega_{q_2} - \omega_{q_1}) = \sum \boldsymbol{\Delta}_{q_4} \mathbf{H}_{q_4} \mathbf{F} \mathbf{R}_{ss} \mathbf{F}^H \mathbf{H}_{q_3}^H \boldsymbol{\Delta}_{q_3}^H \qquad (32)$$

where the sum is over q_4, q_3 that satisfy $\omega_{q_4} - \omega_{q_3} = \omega_{q_2} - \omega_{q_1}$.

It is not difficult to see that we can obtain estimates of $\mathbf{H}_Q$ and $\mathbf{H}_0$ using the fact that $\mathbf{C}_{xx}(\omega_Q - \omega_0) = \boldsymbol{\Delta}_Q \mathbf{H}_Q \mathbf{F} \mathbf{R}_{ss} \mathbf{F}^H \mathbf{H}_0^H \boldsymbol{\Delta}_0^H$, which will enable estimation of $\mathbf{H}_Q$ and $\mathbf{H}_0$ using the method of Section III. Let us now consider $\mathbf{C}_{xx}(\omega_Q - \omega_1)$. If ω_Q and ω_1 are the only pair that gives rise to the difference $\omega_Q - \omega_1$, then we can still use the method of Section III to get $\mathbf{H}_1$. If not, the only other pair that can give rise to this difference is ω_{Q-1} and ω_0. In this case we have

$$\begin{aligned}\mathbf{C}_{xx}(\omega_Q - \omega_1) = {} & \boldsymbol{\Delta}_Q \mathbf{H}_Q \mathbf{F} \mathbf{R}_{ss} \mathbf{F}^H \mathbf{H}_1^H \boldsymbol{\Delta}_1^H \\ & + \boldsymbol{\Delta}_{Q-1} \mathbf{H}_{Q-1} \mathbf{F} \mathbf{R}_{ss} \mathbf{F}^H \mathbf{H}_0^H \boldsymbol{\Delta}_0^H . \qquad (33)\end{aligned}$$

Given $\mathbf{C}_{xx}(\omega_Q - \omega_1)$, $\boldsymbol{\Delta}_Q \mathbf{H}_Q$ and $\boldsymbol{\Delta}_0 \mathbf{H}_0$, suppose we know how to estimate $\mathbf{H}_{Q-1}$ and $\mathbf{H}_1$. Then, according to (32), $\mathbf{C}_{xx}(\omega_Q - \omega_3)$ can contain at most three terms, one of which will be $\boldsymbol{\Delta}_{Q-1} \mathbf{H}_{Q-1} \mathbf{F} \mathbf{R}_{ss} \mathbf{F}^H \mathbf{H}_1^H \boldsymbol{\Delta}_2^H$, which has been estimated, and can be subtracted from $\mathbf{C}_{xx}(\omega_Q - \omega_3)$, which will make the problem identical to solving (33) for $\mathbf{H}_{Q-1}$ and $\mathbf{H}_1$, given $\mathbf{C}_{xx}(\omega_Q - \omega_1)$, $\mathbf{H}_Q$, and $\mathbf{H}_0$. The argument is similar for a general $\mathbf{C}_{xx}(\omega_Q - \omega_q)$ for $q > 3$. Therefore, without loss of generality, to establish identifiability, based on $\mathbf{C}_{xx}(\omega_Q - \omega_1)$, $\boldsymbol{\Delta}_Q \mathbf{H}_Q$, and $\boldsymbol{\Delta}_0 \mathbf{H}_0$, we need to be able to solve for $\mathbf{H}_{Q-1}$ and $\mathbf{H}_1$ in (33). Since we know $\mathbf{H}_0^H$, we also know the L Vandermonde vectors $\mathbf{v}_1, \cdots, \mathbf{v}_L$ in its nullspace. Multiplying (33) by $\mathbf{u}_l := \boldsymbol{\Delta}_0^{-H} \mathbf{v}_l$ and using the structure of $\mathbf{H}_1^H$, we can get rid of the second term in (33) and obtain

$$\begin{aligned}\mathbf{C}_{xx}(\omega_Q - \omega_1)\mathbf{u}_l = {} & \boldsymbol{\Delta}_Q \mathbf{H}_Q \mathbf{F} \mathbf{R}_{ss} \mathbf{F}^H \boldsymbol{\Delta}_1 \mathbf{U}_l \mathbf{h}_1 \\ & l = 1, \cdots, L. \qquad (34)\end{aligned}$$

Because $\mathbf{u}_l$, like $\mathbf{v}_l$, are *also* Vandermonde vectors, $\mathbf{U}_l$ is an $M \times (L+1)$ rank 1 matrix obtained by $\mathbf{U}_l = \mathbf{u}(1:M)\mathbf{u}_l^T(1:L+1)$ Therefore, in order to solve for $\mathbf{h}_1$, we need to use all L equations in (34). Let

$$\mathcal{U} := [(\boldsymbol{\Delta}_Q \mathbf{H}_Q \mathbf{F} \mathbf{R}_{ss} \mathbf{F}^H \mathbf{U}_1)^T \cdots (\boldsymbol{\Delta}_Q \mathbf{H}_Q \mathbf{F} \mathbf{R}_{ss} \mathbf{F}^H \mathbf{U}_L)^T]^T$$

and

$$\mathbf{u} := [(\mathbf{C}_{xx}(\omega_Q - \omega_1)\mathbf{u}_1)^T \cdots (\mathbf{C}_{xx}(\omega_Q - \omega_1)\mathbf{u}_L)^T]^T.$$

Then, the L equations in (34) can be written as

$$\mathcal{U}\,\mathbf{h}_1 = \mathbf{u}. \tag{35}$$

It is not difficult to show that the $PL \times (L+1)$ matrix $\mathcal{U}$ has rank L. This means that (35) can be solved for $\mathbf{h}_1$ up to a scale ambiguity. To obtain $\mathbf{H}_{Q-1}$, one can work with the transpose of $\mathbf{C}_{xx}(\omega_Q - \omega_1)$ and apply the same procedure. Since we are estimating $\mathbf{H}_q$ separately, there is a scalar ambiguity β_q associated with each $\mathbf{H}_q$. Let us assume without loss of generality that $\beta_0 = 1$. Then, we can estimate β_Q as described in Section III. Knowing β_Q, β_1 can be estimated using (34), which shows that even when a2) is not assumed, the scalars $\{\beta_q\}_{q=0}^{Q}$ can be estimated.

APPENDIX B

In this Appendix, we will investigate the relationship between the channel zeros and the rank of $\mathcal{H}$ in (20).

Recall that $\mathcal{H}$ is the filtering matrix arising from the MIMO system whose I/O relationship is given in (19). It was shown in [1] that the following are sufficient conditions for $\mathcal{H}$ to be full column rank.

i) $\mathbf{H}(z)$ is column reduced. This condition is implied by the leading coefficient matrix $\overline{\mathbf{H}}_0 = \Delta_0 \mathbf{H}_0 = \mathbf{H}_0$ being full rank [11, p. 386].

ii) $\mathbf{H}(z) := \sum_{q=0}^{Q} \boldsymbol{\Delta}_q \mathbf{H}_q z^{-q}$ is full rank $\forall z \neq 0$.

Since $\mathbf{H}_0$ is always full rank due to its structure, we need only investigate the rank of $\mathbf{H}(z)$. Let $H_l(z) := \sum_{q=0}^{Q} h_q(l) z^{-q}$. Then, the $P \times M$ polynomial matrix $\mathbf{H}(z)$ is given by

$$\mathbf{H}(z) = \begin{bmatrix} H_0(z) & 0 & \cdots & 0 \\ \vdots & H_0(z\rho) & \ddots & \vdots \\ H_L(z\rho^L) & \vdots & \ddots & 0 \\ 0 & H_L(z\rho^{L+1}) & \ddots & H_0(z\rho^{M-1}) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \vdots & 0 & H_L(z\rho^{P-1}) \end{bmatrix} \tag{36}$$

where $\rho := \exp(-j2\pi/NP)$. This means that it is necessary that $\{H_l(z\rho^l)\}_{l=0}^{L}$ be coprime for $\mathbf{H}(z)$ to be full column rank because without this condition, any one of the columns of $\mathbf{H}(z)$ could be zero, which would make it lose rank.

One should also note that if N is large enough so that $\rho \approx 1$, then the diagonal elements of (36) are equal, in which case, the only way $\mathbf{H}(z)$ could lose rank would be if $\{H_l(z\rho^l)\}_{l=0}^{L}$ have common zeros. Hence, for N large enough, $\{H_l(z\rho^l)\}_{l=0}^{L}$ being coprime suffices for $\mathcal{H}$ to be full rank.

ACKNOWLEDGMENT

The authors would like to thank Dr. A. Scaglione for her comments.

REFERENCES

[1] K. Abed-Meraim, P. Loubaton, and E. Moulines, "A subspace algorithm for certain blind identification problems," *IEEE Trans. Inform. Theory*, vol. 43, pp. 499–511, Mar. 1997.

[2] G. Colman, S. Blostein, and N. Beauleiu, "An ARMA multipath fading simulator," in *Wireless Personal Communications: Improving Capacity, Services and Reliability*. Boston, MA: Kluwer, 1997.

[3] A. V. Dandawate and G. B. Giannakis, "Asymptotic theory of mixed time averages and kth order cyclic-moment and cumulant statistics," *IEEE Trans. Inform. Theory*, vol. 41, pp. 216–232, Jan. 1995.

[4] T. Eyceöz, A. Duel-Hallen, and H. Hallen, "Deterministic channel modeling and long range prediction of fast fading mobile radio channels," *IEEE Commun. Lett.*, vol. 2, pp. 254–256, Sept. 1998.

[5] G. B. Giannakis and C. Tepedelenlioğlu, "Basis expansion models and diversity techniques for blind identification and equalization of time varying channels," *Proc. IEEE*, pp. 1969–1986, Nov. 1998.

[6] ——, "Direct blind equalizers of multiple FIR channels: A deterministic approach," *IEEE Trans. Signal Processing*, vol. 47, pp. 62–74, Jan. 1999.

[7] D. N. Godard, "Self-recovering equalization and carrier-tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. COMM-28, pp. 1867–1875, 1980.

[8] P. Höher, "A statistical discrete-time model for the WSSUS multipath channel," *IEEE Trans. Veh. Technol.*, vol. 41, pp. 461–468, July 1992.

[9] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.

[10] W. G. Jeon, K. H. Chang, and Y. S. Cho, "An equalization technique for orthogonal frequency-division multiplexing systems in time-variant multipath channels," *IEEE Trans. Commun.*, vol. 47, pp. 27–32, Jan. 1999.

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

[12] M. Lamarca and G. Vázquez, "Multichannel receivers for OFDM and TDMA in mobile communications," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, vol. V, Munich, Germany, Apr. 24–24, 1995, pp. 3865–3868.

[13] Y. Li, L. J. Cimini, and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, pp. 902–915, July 1998.

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

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

[16] B. Porat, *Digital Processing of Random Signals*. Englewood Cliffs, NJ: Prentice-Hall, 1994.

[17] T. Rappaport, *Wireless Communications: Principles and Practice*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[18] M. Russell and G. L. Stuber, "Interchannel interference analysis of OFDM in a mobile environment," in *Proc. 45th IEEE Veh. Technol. Conf.*, vol. 2, 1995, pp. 820–824.

[19] A. M. Sayeed, A. Sendonaris, and B. Aazhang, "Multiuser detection in fast-fading multipath environments," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1691–1701, Dec. 1998.

[20] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers—Part II: Blind channel estimation, synchronization and direct equalization," *IEEE Trans. Signal Processing*, vol. 47, pp. 2007–2022, July 1999.

[21] A. Scaglione, S. Barbarossa, and G. B. Giannakis, "Self-recovering equalization of time-selective fading channels using redundant filterbank precoders," in *Proc. DSP Workshop*, Bryce Canyon, UT, Sept. 1998.

[22] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inform. Theory*, vol. 40, pp. 340–349, Mar. 1994.

[23] M. K. Tsatsanis and G. B. Giannakis, "Modeling and equalization of rapidly fading channels," *Int. J. Adapt. Contr. Signal Process.*, vol. 10, pp. 159–176, Mar. 1996.

[24] M. K. Tsatsanis, G. B. Giannakis, and G. Zhou, "Estimation and equalization of fading channels with random coefficients," *Signal Process.*, vol. 53, pp. 211–229, 1996.

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

[26] P. P. Vaidyanathan, *Multirate Systems and Filterbanks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[27] R. A. Ziegler and J. M. Cioffi, "Estimation of time-varying digital radio channels," *IEEE Trans. Veh. Technol.*, vol. 41, pp. 134–151, May 1992.

Cihan Tepedelenlioğlu was born in Ankara, Turkey, in 1973. He received the B.S. degree with highest honors from Florida Institute of Technology, Melbourne, in 1995 and the M.S. degree from the University of Virginia, Charlottesville, in 1998, both in electrical engineering. Since January 1999, he has been working as a research assistant at the University of Minnesota, Minneapolis, and is currently pursuing the Ph.D. degree in the Department of Electrical and Computer Engineering.

His research interests include statistical signal processing, system identification, time-varying systems, and modeling, tracking, and equalization of fading channels in mobile communications.

Georgios B. Giannakis (F'96) received the Diploma in electrical engineering from the National Technical University of Athens, Athens, Greece in 1981. From September 1982 to July 1986, he was with the University of Southern California (USC), Los Angeles, where he received the M.Sc. degree in electrical engineering in 1983, the M.Sc. degree in mathematics in 1986, and the Ph.D. degree in electrical engineering in 1986.

After lecturing for one year at USC, he joined the University of Virginia (UVA), Charlottesville, in 1987, where he became a Professor of electrical engineering in 1997, Graduate Committee Chair, and Director of the Communications, Controls, and Signal Processing Laboratory in 1998. Since January 1999, he has been with the University of Minnesota, Minneapolis, as a Professor of electrical and computer engineering. His general interests lie in the areas of signal processing and communications, estimation and detection theory, time-series analysis, and system identification—subjects on which he published more than 100 journal papers. Specific areas of expertise have included (poly)spectral analysis, wavelets, cyclostationary, and non-Gaussian signal processing with applications to sensor array and image processing. Current research focuses on transmitter and receiver diversity techniques for equalization of single- and multiuser communication channels, mitigation of rapidly fading channels, compensation of nonlinear amplifier effects, redundant filterbank transceivers for block transmissions, multicarrier, and wideband communication systems.

Dr. Giannakis was awarded the School of Engineering and Applied Science Junior Faculty Research Award from UVA in 1988 and the UVA-EE Outstanding Faculty Teaching Award in 1992. He received the IEEE Signal Processing Society's 1992 Paper Award in the Statistical Signal and Array Processing (SSAP) area and the 1999 Young Author Best Paper Award (with M. K. Tsatsanis). He co-organized the 1993 IEEE Signal Processing Workshop on Higher-Order Statistics, the 1996 IEEE Workshop on Statistical Signal and Array Processing, and the first IEEE Signal Processing Workshop on Wireless Communications in 1997. He guest (co-)edited two special issues on high-order statistics (*International Journal of Adaptive Control and Signal Processing* and the EURASIP journal *Signal Processing*) and the January 1997 Special Issue on Signal Processing for Advanced Communications of the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He has served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE SIGNAL PROCESSING LETTERS, a Secretary of the Signal Processing Conference Board, a member of the SP Publications Board, and a member and vice-chair of SSAP Technical Committee. He now chairs the Signal Processing for Communications Technical Committee and serves as Editor-in-Chief of the IEEE SIGNAL PROCESSING LETTERS. He is also a member of the IEEE Fellow Election Committee, the IMS, and the European Association for Signal Processing.