Basis Expansion Models and Diversity Techniques for Blind Identification and Equalization of Time-Varying Channels

GEORGIOS B. GIANNAKIS, FELLOW, IEEE, AND CIHAN TEPEDELENLIOĞLU

Invited Paper

The time-varying impulse response of rapidly fading mobile communication channels is expanded over a basis of complex exponentials that arise due to Doppler effects encountered with multipath propagation. Blind methods are reviewed for estimating the bases' parameters and the model orders. Existing second-order methods are critiqued and novel algorithms are developed for blind identification, direct, zero-forcing equalization and minimum mean square error (MMSE) equalization by combining channel diversity with temporal (fractional sampling) and/or spatial diversity which becomes available with multiple receivers. Illustrative simulations are also presented.

Keywords— *Adaptive equalizers, diversity methods, Doppler effect, fading channels, identification, least mean square methods, mobile communication, time-varying channels.*

I. Introduction

Blind techniques for identification of linear time-invariant (TI) systems have found widespread applications in time series modeling, econometrics, exploration seismology, and equalization of communication channels, just to name a few. With no access to the input, many blind methods have relied on stationary high-order statistics [13], [18], [34], [51] and cyclostationary or multivariate second-order statistics [12], [14], [33], [43], [45], [52] of the output data in order to: 1) either estimate TI systems or 2) their inverses when input recovery is the ultimate goal. Such self-recovering schemes are important, for example, in digital broadcasting because transmission is not interrupted to train new users entering the cell. Similarly, in wireless environments bandwidth is utilized efficiently when cold start-up is possible and in multipoint data networks throughput increases and management overhead drops when training is obviated [18].

Manuscript received August 14, 1997; revised March 6, 1998. This work was supported by NSF Grant MIP 9424305.

The authors are with the Department of Electrical Engineering, University of Virginia, Charlottesville, VA 22903-2442 USA (e-mail: georgios@virginia.edu).

Publisher Item Identifier S 0018-9219(98)06973-4.

However, many systems violate the time-invariance assumption. In cellular telephony, the multipath propagation channel not only exhibits frequency selectivity, which causes intersymbol interference (ISI), but also changes as the mobile communicators move [1], [2], [7], [28], [37], [42]. Temperature and salinity variations cause underwater channels to vary [24], [25], [44], and fluctuations in the ionosphere give rise to deep fades in the data received via microwave links [20], [35]. For channel variations with coherence time in the order of hundreds of symbols (slow fading) adaptive variants of algorithms developed for TI systems offer a valuable alternative, although periodic retraining is recommended to avoid runaway effects [2], [24], [36], [42]. Recursive least-squares (RLS) and least mean-square (LMS) are adaptive algorithms which are known to diverge when channel variations exceed the algorithms' convergence time. In such cases explicit incorporation of the channel's time-varying (TV) characteristics is called for.

Most explicit models of TV communication channels treat the TV taps as uncorrelated stationary random processes which are assumed to be low-pass, Gaussian, with zero mean (Rayleigh fading) or nonzero mean (Rician fading) depending on whether line-of-sight propagation is absent or present [6], [23] [25], [35], [50]. Correlations of the unknown taps capture average channel characteristics and are used to track the channel's time evolution using Kalman filtering estimators [8], [9], [25], [50]. The unobservable channel statistics are either fixed to experimentally computed values [8], [25] or estimated from the data during the decision directed mode [6], [9], [50]. Hidden Markov models have also been used in modeling the tap variations [3].

Statistical modeling is well motivated when TV path delays arise due to a large number of scatterers (e.g., in over-the-horizon communications). But recently deterministic basis expansion models have gained popularity for cellular radio applications, especially when the multipath is caused by a few strong reflectors and path delays exhibit variations due to the kinematics of the mobiles [1], [17], [28, p. 383], [39], [46], [47]. The TV taps are expressed as a superposition of TV bases (e.g., complex exponentials when modeling Doppler effects) with TI coefficients. By assigning time variations to the bases, rapidly fading channels with coherence time as small as a few tens of symbols can be captured. Such finitely parameterized expansions render TV channel estimation tractable and have been previously used in modeling speech and economic time series [21], [29]. They are also encountered in Doppler radar and sonar applications when scintillating point targets give rise to delays which change (linearly or quadratically) with time and cause Doppler shifts in the carrier frequency [38]. In [28, p. 383], it is argued that such Doppler-induced variations are equivalent to the random coefficient model since narrow-band Gaussian processes are well approximated by superimposed sinusoids having constant amplitudes and random phases.

Time- and frequency-selective channels are special cases of the basis expansion models considered here. Although most existing blind equalization research has focused on frequency selective channels, modeling time-selective effects are well motivated due to local oscillator drifts and/or relative motion encountered in mobile communications.

Finite basis expansions offer well-structured parsimonious modeling which allows for blind identification of TV channels. In [47], this important feature was established first based on second- and fourth-order output correlations. The high-variance of high-order TV statistics with moderate data records, prompted recent second-order methods which rely on complementing the TV channel's diversity with time diversity (offered when oversampling the continuous-time output) and/or with spatial diversity (appearing when output data are collected from multiple antennas) [10], [16], [31], [32], [49].

It is the objective of this paper to review, unify, and extend these second-order diversity combining approaches for blind identification and equalization of finite impulse response (FIR) TV communication channels where the variation of the channel is modeled by a basis expansion. To put TV approaches in context, the random model is reviewed briefly in Section II, followed by the basis expansion model introduced in Section III. With the rapidly fading mobile channel as a paradigm, subsequent presentation focuses on cyclostationary methods used to estimate the frequencies of the Fourier bases. Section IV describes blind TV channel estimation methods which utilize the whiteness of the input and rely upon output samples collected at one or two sensors. A deterministic approach is also reviewed along with order selection techniques, which are developed to determine not only the channel memory, but also the number of bases necessary in the expansion (this corresponds to the number of dominant reflectors in a multipath terrain). Mean-square error (MSE, Wiener) equalizers are presented in Section V along with direct blind equalizers derived in a deterministic framework. The latter lend themselves naturally to adaptive schemes and allow almost perfect equalization when the signal-to-noise ratio is high, while imposing minimal assumptions on the input. Representative simulations are given in Section VI, while conclusions, topics not covered, and thoughts for future research are delineated in Section VII (more technical proofs can be found in the Appendixes).

Bold upper (lower) case will denote matrices (column vectors). Prime will stand for Hermittian transpose, * for conjugate, ^{T} for transpose, $\dagger$ for pseudo-inverse, $\otimes$ for Kronecker product, $\mathcal{R}$ for range, and $\mathcal{N}$ for null space.

II. Fading Channels: Random Models

In some communication schemes, unpredictable changes in the medium warrant modeling the TV impulse response (TVIR) $h_c(t;\tau)$ as a stochastic process in the time variable t. Using central limit theorem arguments, the TVIR is usually approximated as a complex Gaussian process. It is common practice to assume the channel to be wide sense stationary for a fixed lag τ and uncorrelated for different lags (i.e., wide sense stationary uncorrelated channel (WSSUC) assumption) [35]. The channel spectral density for a fixed τ is called the scattering function $S(\omega;\tau)$, and it fully characterizes the second-order statistics of the WSSUC. There are several other functions that rely on the second-order statistics of the random channel. The integral $S_C(\omega) := \int S(\omega;\tau)\,d\tau$ is called the Doppler spectrum, and its extent, the Doppler spread, is a measure of the channel's time variation. The so-called multipath intensity profile $\phi(\tau) := E|h_c(t;\tau)|^2$ describes how the output power varies as a function of the delay τ, and the length of its support is called the multipath spread and is a measure of the average extent of the multipath.

The characterization of the random channel mainly has been used to analyze and simulate existing methods rather than to undo the TV distortion the fading channel has on the input signal. However, recent work in [6], [25], and [50] has addressed the TV channel identification problem by casting it in discrete time and using a Kalman filter to track the channel parameters.

Consider the fading communication system model of Fig. 1 before the sampler with the input/output (I/O) relationship

$$x_c(t) = \sum_{l=-\infty}^{\infty} s(l) h_c(t; t - lT_s) + v_c(t) \qquad (1)$$

where subscript $_c$ denotes continuous time, $h_c(t;T)$ is the convolution of the spectral pulse $f_c^{(tr)}(t)$, the TV impulse response $f_c^{(ch)}(t;\tau)$, the receive-filter $f_c^{(rec)}(t)$, $s(l)$ is the sequence of input symbols, and $v_c(t)$ is the noise process. If the output $x_c(t)$ is sampled at the symbol rate $1/T_s$, Fig. 1 can be simplified into Fig. 2 with an I/O relation in discrete-time

$$x(n) = \sum_{l=0}^{L} h(n;l)s(n-l) + v(n), \quad n = 0, 1, \cdots, N-1 \qquad (2)$$

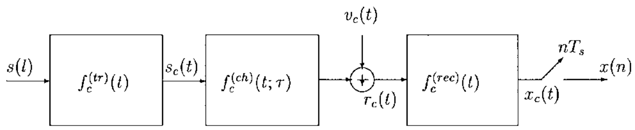

Fig. 1. Continuous-time TV communication system.

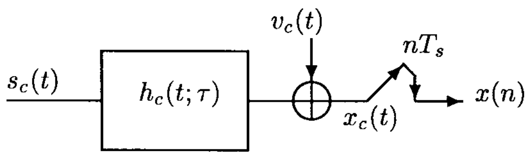

Fig. 2. TV channel model.

where $h(n;\, l) := h_c(nT_s;\, lT_s)$ is truncated to an order L, which is common practice in communications applications. Note that if $h(n;\, l) = h(l) \,\forall n$, then (2) yields a time invariant frequency selective channel and the I/O relationship becomes $x(n) = \sum_{l=0}^{L} h(l)s(n-l) + v(n)$. On the other hand, if $L = 0$, (2) yields a time selective channel with an I/O relationship, $x(n) = h(n)s(n) + v(n)$. One approach in characterizing the variation of the impulse response $h(n;\, l)$ is to consider it as a stochastic process in the time index n. In communications, tracking the variations of the channel taps is of importance [35]. To this end, fitting parametric models to the variation of the channel coefficients has been proposed in [25] and [50]. In [50], the challenging task of estimating random channel parameters from I/O data $\{s(n),\, x(n)\}_{n=0}^{N-1}$ has been tackled. The channel was assumed to obey an $(L+1) \times 1$ vector AR(p) model

$$\mathbf{h}(n) = \sum_{i=1}^{p} \mathbf{A}(i)\mathbf{h}(n-i) + \mathbf{u}(n) \tag{3}$$

where $\mathbf{h}(n) := [h(n;\, 0) \cdots h(n;\, L)]^T$, and $\mathbf{u}(n)$ is an independently, identically distributed (i.i.d) circular complex Gaussian vector process whose components are uncorrelated with each other. The coefficient matrices $\{\mathbf{A}(i)\}_{i=1}^{p}$ were estimated using the multichannel Yule–Walker equations

$$\mathbf{R}_{hh}(\tau) = \sum_{i=1}^{p} \mathbf{A}(i)\mathbf{R}_{hh}(\tau - i) + \sigma_u^2 \delta(\tau)\mathbf{I}, \quad \tau = 0, \cdots, p \tag{4}$$

where $\mathbf{R}_{hh}(\tau) := E[\mathbf{h}(n)\mathbf{h}'(n+\tau)]$ is the channel correlation matrix whose entries were estimated consistently from output statistics conditioned on the input [50]. With $\hat{\mathbf{R}}_{hh}(\tau)$ available, we can solve for $\hat{\mathbf{A}}(i)$ using (4). Once the AR parameter matrices are estimated, a Kalman filter is employed to track the channel coefficients after casting the AR model in (3) in a state-space form

$$\begin{aligned} \mathbf{h}_{ss}(n+1) &= \mathcal{A}\mathbf{h}_{ss} + \mathcal{J}\mathbf{u}(n) \\ x(n) &= [s(n) \cdots s(n-L)]\mathbf{h}_{ss}(n) + v(n) \end{aligned} \tag{5}$$

where $\mathbf{h}_{ss}(n) := [\mathbf{h}^T(n) \cdots \mathbf{h}^T(n-p+1)]^T$ is the channel state vector, $\mathcal{A}$ is a constant matrix consisting of the AR parameter matrices $\{\mathbf{A}(i)\}_{i=1}^{p}$ in its first block row, $p-1$ identity matrices in its first sub-block diagonal and zero elsewhere, and $\mathcal{J} := [\mathbf{I}\, \mathbf{0} \cdots \mathbf{0}]$. Finally, a decision feedback equalizer using the channel estimates is utilized to obtain an estimate of the input (see [50] and references therein).

III. Fading Channels: Basis Expansion Models

Consider the random variation in one tap, l_0, of a multipath mobile radio channel [27], [28]

$$h(n;\, l_0) = \sum_{q=1}^{Q} c_q(l_0) e^{j\phi_q} e^{j(2\pi v/\lambda)\, \cos(2\pi q/Q)n} \tag{6}$$

where c_q is the amplitude of the qth path, ϕ_q is a uniformly distributed random variable in $[0,\, 2\pi]$, λ is the wavelength corresponding to the carrier frequency, and v is the speed of the mobile [28, p. 382]. For sufficiently large Q, the amplitude of (6) approximates a Rayleigh probability density function (pdf), and the power spectrum of (6) provides a discrete approximation to experimentally measured fading spectra which are of the form $S(\omega;\, l_0) = A(l_0)[1 - (\omega - \omega_c)^2\lambda^2/v^2]^{-1/2}$, where $A(l_0)$ is a constant determining the power of tap l_0, and ω_c is the carrier frequency [27], [28].

As an alternative to the random channel assumption of the previous section, where $h(n;\, l)$ is a realization of a stochastic process, the variation in the impulse response can be captured deterministically by means of a basis expansion

$$h(n;\, l) = \sum_{q=1}^{Q} \overline{h}_q(l) b_q(n) \tag{7}$$

where the TI parameters $\{\overline{h}_q(l)\}_{q=1}^{Q}$, together with the bases $\{b_q(n)\}_{q=1}^{Q}$ characterize the system. It is clear that (7) with $\overline{h}_q(l) := c_q(l)e^{j\phi_q}$ and $b_q(n) := \exp(j\omega_q n)$ with $\omega_q := (2\pi v/\lambda)\, \cos(2\pi q/Q)$ have the same functional form as the model in (6). In Fig. 3, we depict how time selectivity, frequency selectivity, and time-frequency selectivity manifest themselves in plots generated by (6), which is subsumed by the basis expansion model discussed in this section. As the number of paths Q increases (chosen to be ten in Fig. 3), the basis expansion model approximates the well-known random coefficient fading models used to simulate mobile communication channels [27], [28].

In summary, random coefficient models are used either for identification of the model parameters, which determine the evolution of the channel coefficients, from the stationary moments of the output, or they are used for simulating fading channels with certain spectral properties. Interestingly, random coefficient models used to simulate mobile channels can be obtained from the basis expansion model with random parameters. In this paper we will focus on terrains entailing only a few reflectors so that the Doppler and multipath parameters can be considered deterministic. We will rely on the basis expansion channel model to perform blind identification and equalization.

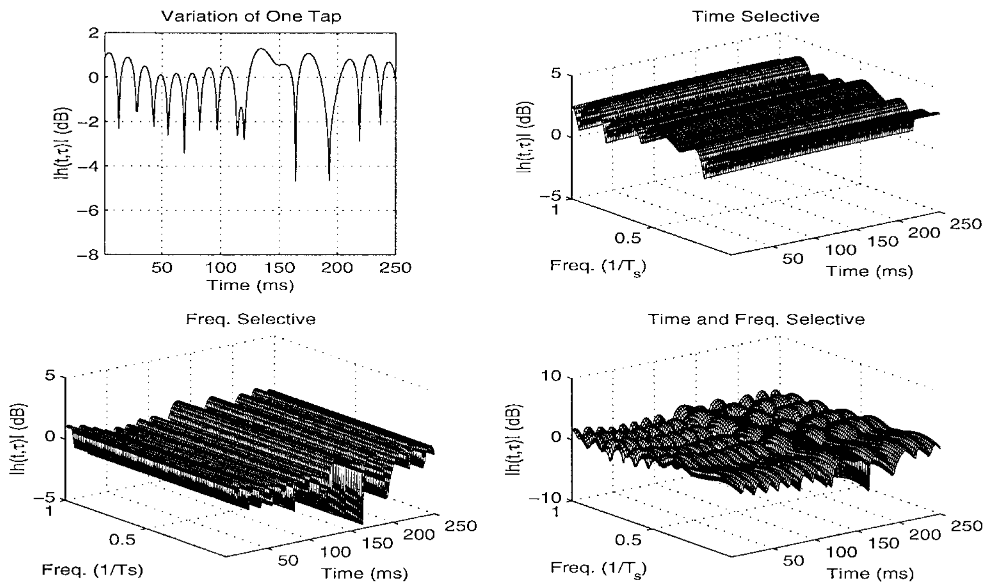

Fig. 3. Fading channels generated by the basis expansion model.

A. Exponential Basis Expansion Model

To appreciate the usefulness of complex exponential bases, consider a communication signal

$$s_c(t) = \mathrm{Re}\left\{ e^{j\omega_c t} \sum_l s(l) f_c^{(tr)}(t - lT_s) \right\}$$

transmitted through a TV multipath channel

$$f_c^{(ch)}(t;\tau) = \sum_{q=1}^{Q} A_q(t)\delta(\tau - d_q(t))$$

where $s(l)$ are the input symbols, Q is the number of paths, and $A_q(t)$, $d_q(t)$ denote each path's TV attenuation and delay, respectively. With reference to Fig. 1, we convolve $s_c(t)$ with $f_c^{(ch)}(t;\tau)$ and remove the carrier to arrive at the received signal-plus-noise model in baseband form: $r_c(t) = \sum_{q=1}^{Q} A_q(t)s_c(t - d_q(t)) + v_c(t)$. To suppress the additive white noise Gaussian (AWGN) $v_c(t)$, we filter $r_c(t)$ through the receive-filter $f_c^{(rec)}(t)$ and obtain

$$\begin{aligned} x_c(t) &= \sum_l \Bigg\{ s(l) \Bigg[\sum_{q=1}^{Q} \int_{(l-1)T_s}^{lT_s} A_q(\tau) f_c^{(tr)}(\tau - lT_s - d_q(\tau)) \\ &\qquad \cdot f_c^{(rec)}(t-\tau) e^{j\omega_c d_q(\tau)}\, d\tau \Bigg] \\ &\qquad + \int_{(l-1)T_s}^{lT_s} v_c(\tau) \cdot f_c^{(rec)}(t-\tau)\, d\tau \Bigg\}. \end{aligned} \quad (8)$$

Let $f_2(t) := \int_{T_s} f_c^{(tr)}(\tau) f_c^{(rec)}(t-\tau)\, d\tau$ denote the time invariant (TI) transmit-receive filters in cascade, and assume the following:

a1) constant attenuation and delay over a symbol, i.e., $A_q(\tau) = \text{const.} := A_q(l)$, for $\tau \in [(l-1)T_s, lT_s]$, $d_q(\tau) = \text{const.} := d_q(l)$, for $\tau \in [(l-1)T_s, lT_s]$;

a2) linearly varying delays across symbols (valid for approximately constant path velocity), i.e., $d_q(l) = \nu_q l + \epsilon_q$, where ν_q is proportional to the path velocity and T_s. This is a first-order approximation of the delay variation. Existence of higher order terms would yield polynomial phase signals, which brings up the tradeoff between accuracy and complexity; this is outside the scope of this paper.

Under a1), we can pull $A_q(l)$ and $d_q(l)$ outside the integral in (8), using the definition of $f_2(t)$ and after a change of variables we have $x_c(t) = \sum_l s(l)h_c(t; t - lT_s) + v_c(t)$, where

$$h_c(t; t - lT_s) := \sum_{q=1}^{Q} A_q(l) f_2(t - lT_s - d_q(l)) e^{j\omega_c d_q(l)}. \quad (9)$$

After sampling the output at the symbol rate $1/T_s$ (fractional sampling will be considered in Section IV-B), and using a2), we obtain $h_c(nT_s; lT_s) = \sum_{q=1}^{Q} A_q(n - l) f_2(nT_s - \nu_q(n-l) - \epsilon_q) \exp[j\omega_c \nu_q(n-l) + \epsilon_q]$. If we further assume that the $h_q(l) := A_q(n-l) f_2(nT_s - \nu_q(n-l) - \epsilon_q) \exp(j\omega_c \epsilon_q)$ is approximately constant with respect to n since it is changing slowly compared to the exponential, we obtain

$$h(n; l) = \sum_{q=1}^{Q} h_q(l) e^{j\omega_q(n-l)} \quad (10)$$

with $h(n;l) := h_c(nT_s; lT_s)$ and $\omega_q := \omega_c \nu_q$. Notice that with $\overline{h}_q(l) := h_q(l) \exp(-j\omega_q l)$ we arrive at the basis expansion model in (7); so the exponentials' dependence on lag l can be included in the parameters, yielding an I/O relationship (see also Fig. 4)

$$\begin{aligned} x(n) &= \sum_{q=1}^{Q} \left[\sum_{l=0}^{L} \overline{h}_q(l) b_q(n) s(n-l) \right] + v(n) \\ b_q(n) &= e^{j\omega_q n}. \end{aligned} \quad (11)$$

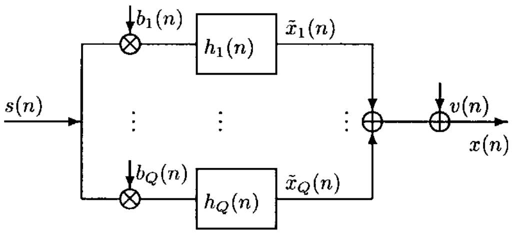

Fig. 4. Multichannel discrete-time equivalent of a TV basis expansion model.

The complex exponentials in (11) can be viewed as each path's Doppler arising due to motion—an effect also encountered in radar and sonar where moving targets induce TV delays which for narrow-band signals manifest themselves as TV phases [38]. Since the same input is modulated by Q different complex exponentials in Fig. 4, some redundancy is introduced at the output which we call channel (or Doppler) diversity, a term also used in [39] in a time-frequency context.

Our TV channel parametrization is not unique. An alternative one is proposed in [40] as

$$x(n) = \sum_{l=0}^{\hat{L}} \hat{h}(l)s(n-\tau_l)e^{j2\pi f_l n} + v(n) \qquad (12)$$

where τ_l is the delay (expressed in multiples of the symbol interval T_s), and f_l denotes the Doppler frequency shift (normalized by the symbol rate $1/T_s$) of the lth path relative to the zeroth path. The I/O relation in (12) is simpler since it only involves a single sum, but it can be shown that the basis expansion model can capture a more general class of time variations, and hence (11) subsumes (12). In view of this fact, we focus on the basis expansion model in (11).

In this paper, given $x(n)$, $n = 0, \cdots, N-1$, we would like to do the following : 1) estimate $\{\omega_q\}_{q=1}^Q$; 2) determine the channel length L and the number of bases Q; and 3) estimate $\{h_q(l)\}_{q=1}^Q$, or the equalizers $g(n;k)$ which, when convolved with the data $x(n)$, yield input estimates $\hat{s}(n)$, and estimate the equalizer length K. First, we will address frequency estimation.

B. Estimating the Exponential Bases

In all the methods that follow, to estimate the TI parameters $\{h_q(l)\}$, we will assume the knowledge of the bases. So for complex exponential bases, the question of estimating the frequencies $\{\omega_q\}_{q=1}^Q$ from the output $x(n)$ in (11) needs to be addressed.

The idea is to exploit the cyclostationarity of $x(n)$ and use its TV moments that only depend on the time index n through the complex exponentials [46]. The frequencies of these exponentials are calculated from the so-called cyclic moments, the Fourier series of the TV moments (e.g., see [12] for detailed definitions). The input $s(n)$ will be assumed to be independent of the noise $v(n)$, white, with mean μ_s and variance σ_s^2. Let

$$\begin{aligned} m_{kl_x}(n;\tau_1,\cdots,\tau_{k+l-1}) :=& E[x^*(n)\cdots x^*(n+\tau_{k-1}) \\ & \cdot x(n+\tau_k)\cdots x(n+\tau_{k+l-1})] \end{aligned} \qquad (13)$$

where $k(l)$ is the number of conjugated (unconjugated) terms, so that, e.g., $m_{01_x}(n) = E[x(n)]$, and $m_{11_x}(n;\tau) = E[x(n)x^*(n+\tau)]$. In this notation, the dependence on n will be dropped when the process is stationary. Clearly, if the input has nonzero mean, the frequencies can be found by computing the Fourier Series of $E[x(n)] = \mu_s \sum_{q=1}^Q [\sum_{l=0}^L h_q(l)] \exp(j\omega_q n) + \mu_v$. A simple calculation on (11) will reveal that if any moment of the input is zero, then the corresponding moment of the output will also be zero, thereby preventing us from estimating $\{\omega_q\}_{q=1}^Q$. If the input is coming from a real, zero-mean constellation such as binary PAM, then $m_{02_s}(0) \neq 0$ and

$$\begin{aligned} m_{02_x}(n;0) =& m_{02_s}(0) \sum_{q_1,q_2=1}^Q \left[\sum_{l=0}^L h_{q_1}(l)h_{q_2}(l)\right] \\ & \cdot e^{j(\omega_{q_1}+\omega_{q_2})n} + m_{02_v}(0) \end{aligned} \qquad (14)$$

enables us to find the frequencies $\Omega_{q_1+q_2} := \{\omega_{q_1} + \omega_{q_2}\colon q_1, q_2 = 1, \cdots, Q\}$, since the only dependence of (14) on n is through the exponentials. The zero lag is chosen for convenience, but if the term in the brackets in (14) is small for some pair q_1, q_2, then different lags $\tau \neq 0$ could be utilized. From $\Omega_{q_1+q_2}$ it is possible to obtain $\{\omega_q\}_{q=1}^Q$ as follows: let $\omega_1 < \cdots < \omega_Q$. Then, $\min(\Omega_{q_1+q_2}) = 2\omega_1$ from which we can find ω_1. The next smallest frequency is $\omega_1 + \omega_2$, from which ω_2 could be found. Knowing ω_2, we can discard $2\omega_2$ from $\Omega_{q_1+q_2}$, since we do not know whether $\omega_1 + \omega_3 > 2\omega_2$, and find ω_3 from $\omega_1 + \omega_3$. This procedure enables the computation $\{\omega_q\}_{q=1}^Q$ from the knowledge of $\Omega_{q_1+q_2}$.

Unfortunately, for a class of important constellations (4-quadrature amplitude modulation (QAM), 16-QAM), due to their symmetry, the unconjugated correlation of the input $m_{02_s}(0) \equiv 0$, therefore $m_{02_x}(n;0) \equiv 0$ when the symbols are equiprobable. Thus, we are prompted to use

$$m_{11_x}(n;0) = \sigma_s^2 \sum_{q=1}^Q \sum_{l=0}^L h_{q_1}(l)h_{q_2}^*(l)e^{j(\omega q_1 - \omega q_2)n} + \sigma_v^2$$

which enables the estimation of $\Omega_{q_1-q_2} := \{\omega_{q_1} - \omega_{q_2}\colon q_1, q_2 = 1, \cdots, Q\}$. But it is not possible to obtain $\{\omega_q\}_{q=1}^Q$ from $\Omega_{q_1-q_2}$, so higher order moments of the output must be used. Due to their symmetry, all odd ordered moments of many constellations are zero, but their fourth-order moments are nonzero. It is possible to obtain $\Omega_{q_1+q_2+q_3+q_4} := \{\omega_{q_1} + \omega_{q_2} + \omega_{q_3} + \omega_{q_4}\colon q_1, q_2, q_3, q_4 = 1, \cdots, Q\}$ from $m_{04_x}(n;\mathbf{0})$. The method discussed earlier to obtain $\{\omega_q\}_{q=1}^Q$ from $\Omega_{q_1+q_2}$ can also be employed, with slight modifications, to calculate the frequencies from $\Omega_{q_1+q_2+q_3+q_4}$. Since second-order statistics generally have lower variance than higher order statistics, knowledge of the difference frequencies $\Omega_{q_1-q_2}$, whose estimates rely on second order statistics, can be incorporated in the above procedure [46].

The frequencies can be obtained using sample estimates of the cyclic moments which are defined as the Fourier

Series of (13) with respect to n. The estimators for the cyclic moments are [12]

$$\hat{M}_{kl_x}(\alpha;\tau_1,\cdots,\tau_{k+l-1}) := \frac{1}{N}\sum_{n=0}^{N-1} x^*(n)\cdots x^*(n+\tau_{k-1}) \cdot x(n+\tau_k)\cdots x(n+\tau_{k+l-1}) \cdot e^{-j\alpha n} \qquad (15)$$

which can be computed efficiently by taking the fast Fourier transform (FFT) of the output product. For example, in the case of a pulse amplitude modulation (PAM) constellation, as explained earlier in this section, the smallest frequency can be obtained by $\hat{\omega}_1 = (1/2)\arg\min_\alpha |\hat{M}_{02_x}(\alpha;\tau)|$. Cyclic moment estimators are known to be asymptotically normal and mean square consistent when the input has finite moments and the subchannels are of finite length, so that the output satisfies the necessary mixing conditions [5], [12].

IV. Blind TV Channel Identification

Throughout the rest of the paper, the channel coefficients $h_q(l)$ will be assumed to be deterministic, but some approaches (which we term "statistical") will require the input $s(n)$ to be random and white. The bases $\{b_q(n)\}_{q=1}^Q$ are assumed to be known.

A. Statistical Approach 1: One Sensor

Here we will not necessarily assume the bases are complex exponentials, for reasons that will soon be given. From the correlations of $x(n)$ in (11), we can obtain

$$m_{11_x}(n;\tau) = \sum_{q_1,q_2=1}^{Q}\left[\sigma_s^2 \sum_{l=0}^{L} h_{q_1}(l)h_{q_2}^*(l+\tau)\right] b_{q_1}(n) \cdot b_{q_2}^*(n+\tau). \qquad (16)$$

Given $m_{11_x}(n;\tau)$, if the product sequences $\{b_{q_1}(n)b_{q_2}^*(n+\tau), q_1, q_2 = 1,\cdots,Q\}$ are linearly independent, by solving the linear equations in (16) we can obtain the (deterministic) correlations of all possible TI coefficient pairs of channels

$$r_{q_1q_2}(\tau) := \sum_{l=0}^{L} h_{q_1}(l)h_{q_2}^*(l+\tau), \qquad q_1, q_2 = 1,\cdots,Q. \qquad (17)$$

The problem of obtaining $\{h_q(l)\}_{q=1}^Q$ from (17) can be solved with conventional subspace approaches if $r_{q_1q_2}(\tau)$ has been estimated from (16). Subspace approaches have been used to estimate a set of coprime TI channels excited by a common white input $w(n)$ (see Fig. 5) [33], [45]. Notice that the output auto and cross correlations of the single-input multiple-output (SIMO) system in Fig. 5 provides (up to a scale) all possible deterministic correlations in (17) of a set of FIR channels $\{h_q(l)\}_{q=1}^Q$ to be estimated. Hence, both the problem of estimating $\{h_q(l)\}_{q=1}^Q$ from $r_{q_1q_2}(\tau)$, as encountered in the TI SIMO blind identification problem, and the TV single-input single-output (SISO) problem of this section can be solved by the same subspace method.

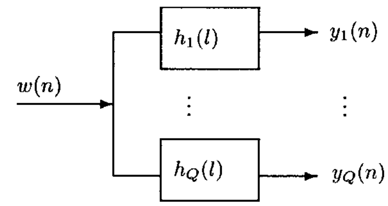

Fig. 5. TI SIMO model.

Theorem 1 [49]: Sufficient conditions for identifiability of $\{h_q(l)\}_{q=1}^Q$ from $m_{11_x}(n;\tau)$ are: 1) for every fixed τ the product sequences $b_{q_1}(n)b_{q_2}^*(n+\tau), q_1, q_2 = 1,\cdots,Q$ are linearly independent; 2) the polynomials $H_q(z) := \sum_{l=0}^L h_q(l)z^{-l}, q = 1,\cdots,Q$, do not have common roots; and 3) the product sequences $b_{q_1}(n)b_{q_2}^*(n+\tau)$ are bounded $\forall\ \tau$, and $\lim_{N\to\infty} 1/N \sum_{n=0}^{N-1}[\mathbf{b}(n)\mathbf{b}'(n)\otimes\mathbf{b}^*(n+\tau)\mathbf{b}^T(n+\tau)]$ is invertible, where $\mathbf{b}(n) := [b_1(n)\cdots b_Q(n)]^T$. □

The method developed in [49] entails two steps: first $r_{q_1q_2}(\tau)$ is obtained from $m_{11_x}(n;\tau)$, where we need 1), since a matrix whose columns are formed by $b_{q_1}(n)b_{q_2}^*(n+\tau)$ needs to be inverted. Second, $\{h_q(l)\}_{q=1}^Q$ needs to be recovered from $r_{q_1q_2}(\tau)$, where 2) becomes necessary. The nontrivial task of estimating the TV statistics $\hat{m}_{11_x}(n;\tau)$ is handled by using an instantaneous estimate $x(n)x^*(n+\tau)$, and consistency of $\hat{r}_{q_1q_2}(\tau)$ is established under 3), which requires some additional boundedness conditions on the bases, and 1) to hold in the limit.

The problem with the method derived from Theorem 1 is that the linear independence assumption a2) on the bases is often not satisfied in practice. Nevertheless, the single sensor approach illustrates nicely that TV channels offer diversity not available with TI channels, and from this point of view blind identification based on second-order statistics is easier for TV channels of structured variation than TI channels. However, the linear independence condition necessary with a single sensor does not hold for complex exponentials, since for $\tau = 0$, $b_1(n)b_1^*(n) = b_2(n)b_2^*(n) = 1\,\forall n$. This brings about looking for alternative ways of obtaining complementary diversity.

B. Statistical Approach 2: Two Sensors

Just like the TI case [15], [52], sampling faster than the symbol rate creates diversity that enables the problem to be cast into a SIMO framework. Suppose $x_c(t)$ in (1) is sampled at a rate M/T_s, where $h_c(t;t-lT_s)$ is given in (9). We obtain the discrete time model $x(n) := x_c(nT_s)/M) = \sum_l s(l)h(n;n-lM)+v(n)$, where $h(n;l) := h_c(nT_s/M;lT_s/M)$ and $v(n) := v_c(nT_s/M)$. Oversampling offers diversity manifested in the M subprocesses defined as $\{x^{(m)}(n) := x(nM+m-1)\}_{m=1}^M$. In the filterbank literature, $x^{(m)}(n)$ are termed the polyphase components of $x(n)$ and can be expressed in terms of the M subchannels $h^{(m)}(n;l)$ and the corresponding noise $v^{(m)}(n)$ as

$$x^{(m)}(n) = \sum_{l=0}^{L} h^{(m)}(n; l)s(n-l) + v^{(m)}(n) \qquad m = 1, \cdots, M \tag{18}$$

where $h^{(m)}(n;l){:=}h(nM+m-1;lM+m-1){:=}h_c(T_s(nM+m-1)/M;T_s(lM+m-1)/M) = \sum_{q=1}^{Q} A_q(n-l)f_2(T_s(lM+m-1)/M - d_q(n-l))\exp(j\omega_c d_q(n-l))$. Because the variation of A_q and f_2 with respect to n is often negligible relative to that of the exponential, it is reasonable to assume the following:

a3.1) $A_q(n-l) \approx \overline{A}_q(l)$;
a3.2) $f_2(T_s(lM+m-1)/M - d_q(n-l)) \approx \overline{f}_2(T_s(lM+m-1)/M - d_q(l))$.

Based on a1)–a3) we have

$$h^{(m)}(n; l) := \sum_{q=1}^{Q} h_q^{(m)}(l)b_q(n-l), \qquad b_q(n) := e^{j\omega_c \nu_q n} \tag{19}$$

where $h_q^{(m)}(l) := \overline{A}_q(l)\overline{f}_2(T_s(lM+m-1)/M + \nu_q l - \epsilon_q)\exp(j\omega_c\epsilon_q)$. Combining (18) and (19) and stacking the M-channel data $\mathbf{x}(n) := [x^{(1)}(n) \cdots x^{(M)}(n)]^T$ and $\mathbf{h}_q(l) := [h^{(1)}(l) \cdots h^{(M)}(l)]^T$ we obtain

$$\mathbf{x}^T(n) = \sum_{q=1}^{Q} \left[\sum_{l=0}^{L} \mathbf{h}_q^T(l)e^{j\omega_q n}s(n-l)\right] + \mathbf{v}^T(n) \tag{20}$$

where, as in (11), the exponentials' dependence on l is absorbed in the TI vector impulse response.

Multichannel diversity can also be achieved by using multiple antennas at the receiver [31], the number of which will be denoted also by M. With the availability of oversampling (time diversity) or multiple sensors (space diversity) the following question arises: what is $M_{\min}$, the minimum M in order to guarantee identifiability without restrictions on the frequencies of the exponential bases? As we will see in Section IV-C, zero-forcing FIR solutions require $M_{\min}$ to be on the order of Q. If the input can be assumed to be white and random, on the other hand, $M_{\min} = 2$, which does not depend on Q and motivates the two-sensor approach of this section.

Consider the I/O relation in (20). Given $x^{(m)}(n)$, $m = 1, \cdots, M$, $n = 0, \cdots, N-1$, a white input sequence $s(n)$, and a distinct set of Q cycles $\omega_1 < \omega_2 < \cdots, < \omega_Q$, the goal is to identify $\{h_q^{(m)}(l), q = 1, \cdots, Q, m = 1, 2\}$. Since the input is white, the output correlations of the two channels are

$$\begin{aligned} &E\Big[x^{(m_1)}(n)x^{*(m_2)}(n+\tau)\Big] \\ &\quad = \sigma_s^2 \sum_{q_1, q_2=1}^{Q} e^{j(\omega_{q_1}-\omega_{q_2})n} \sum_{l=0}^{L} h_{q_1}^{(m_1)}(l-\tau)h_{q_2}^{*(m_2)}(l) \\ &\qquad \cdot e^{-j\omega_{q_2}\tau} + m_{11_v}(\tau). \end{aligned} \tag{21}$$

The Fourier Series coefficients of the (almost) periodic sequence of n in (21) are

$$\begin{aligned} &C_{xx}^{(m_1, m_2)}(\alpha; \tau) \\ &\quad = \sigma_s^2 \sum_{q_1, q_2=1}^{Q} \delta(\alpha - \omega_{q_1} + \omega_{q_2}) \sum_{l=0}^{L} h_{q_1}^{(m_1)}(l-\tau)h_{q_2}^{*(m_2)} \\ &\qquad \cdot (l)e^{-j\omega_{q_2}\tau} + \delta(\alpha)m_{11_v}(\tau). \end{aligned} \tag{22}$$

Taking the z-transform of (22) with respect to τ and assuming that $\alpha \neq 0$, we arrive at the so-called cross-cyclic spectrum

$$\begin{aligned} S_{xx}^{(m_1, m_2)}(\alpha; z) = &\sigma_s^2 \sum_{q_1, q_2=1}^{Q} \delta(\alpha - \omega_{q_1} + \omega_{q_2})H_{q_2}^{*(m_2)} \\ &\cdot (z^* e^{-j\omega_{q_2}})H_{q_1}^{(m_1)}(e^{-j\omega_{q_2}}/z), \\ &\alpha \neq 0. \end{aligned} \tag{23}$$

Identification of the TI subchannels $\{h_q^{(m)}(l)\}_{m=1}^{M}$ is achieved by choosing the appropriate cycles α in the cyclic spectra of (23) so that only a few unknown terms out of the summation survive. In the set $\Omega_{q_1-q_2}$, defined in Section III-B, at least one difference (namely, $\alpha = \omega_Q - \omega_1$) lets a single term survive out of (23), which is the product $H_Q^{*(m_2)}(z^* e^{-j\omega_Q})H_1^{(m_1)}(e^{-j\omega_Q}/z)$. In Appendix I it is shown that this product enables estimation of the subchannels corresponding to the minimum and the maximum frequencies: $\{h_1^{(m)}(l), h_Q^{(m)}(l), m = 1, \cdots, M\}$. After estimating all subchannels corresponding to frequencies in $\Omega_{q_1-q_2}$, that force all but one term in (23) to be zero, it can be shown that there is a way to use (23) by choosing α from $\Omega_{q_1-q_2}$ in decreasing order so that the sum will only contain two products that have unknown subchannels in it. This is all summarized in Theorem 2 (see Appendix I for a proof).

Theorem 2: For $M = M_{\min} = 2$, $b_q(n) = \exp(j\omega_q n)$, and any set of frequencies $\omega_1 < \cdots < \omega_Q$, the following is possible.

1) To identify the subchannels $h_{q_1}^{(m)}(l)$ for q_1, such that there exists a q_2 with $\omega_{q_2} - \omega_{q_1} = \omega_{q_4} - \omega_{q_3} \Rightarrow \omega_{q_2} = \omega_{q_4}$ [in other words, subchannels $h_{q_1}^{(m)}(l)$ for q_1 such that there exists a q_2 that enables only one product to survive out of (23)] with the choice $\omega = \omega_{q_2} - \omega_{q_1}$. The identifiability condition for these subchannels is that $\{H_q^{(m)}(z), m = 1, 2\}$ are coprime for all $q \in [1, Q]$.
2) After estimating all subchannels characterized in 1) (among which are $\{h_Q^{(m)}(l), h_1^{(m)}(l), m = 1, 2\}$), to identify the remaining subchannels using $\{S_{xx}^{(m_1, m_2)}(\omega_Q - \omega_p; z)\}_{p=2}^{Q-1} \cup \{S_{xx}^{(m_1, m_2)}(\omega_p - \omega_1; z)\}_{p=2}^{Q-1}$. Identifiability is guaranteed if $H_Q^{*(m_2)}(z^* e^{-j\omega_Q})$ and $H_1^{(m_1)}(e^{-j\omega_{q_1}}/z)$ are coprime for $m_1, m_2 = 1, 2$, whenever there exists a q_2 with $\omega_Q - \omega_{q_2} = \omega_{q_1} - \omega_1$. $\square$

Multichannel diversity removes the severe conditions on the basis functions from which [49] suffers. In addition to allowing the minimum diversity ($M_{\min}$) for bases of arbitrary frequencies, the two-cycle method also identifies the channel coefficients by use of the cyclic correlations that avoid the zero cycle ($\omega_{q_1} - \omega_{q_2} \neq 0$). This makes additive stationary noise tolerable down to low SNR's [12], a feature also illustrated in the simulations.

C. Indirect Deterministic Approach

In this section we will show how, with sufficient diversity, it is possible to estimate the subchannels and obtain perfect estimates of the input in the absence of noise. Similar to the approaches in [15] and [52] for TI systems, these (so called "deterministic") methods do not require the input to be white or random, thereby allowing the use of coded inputs. Unlike the statistical approaches, reliable identification will be possible with short data records if the SNR is high enough.

First, we will discuss a subspace approach that we term "indirect approach," introduced in [31] and [32]. In Section V-B, direct blind equalizers will be derived under almost identical assumptions.

In order to cast (20) in matrix form, we let

$$\mathbf{s}_q(n) := \left[e^{j\omega_q n} s(n) \cdots e^{j\omega_q (n-L-K)} s(n-L-K)\right]^T$$
$$\mathbf{h}_q := \left[\mathbf{h}_q^T(0) \cdots \mathbf{h}_q^T(L)\right]^T$$

and define for each $q \in [1, Q]$ the $(L+K+1) \times M(K+1)$ block Toeplitz matrix

$$\mathbf{H}_q = \mathcal{T}(\mathbf{h}_q) := \begin{bmatrix} \mathbf{h}_q^T(0) & \cdots & \mathbf{0}^T \\ \vdots & \ddots & \vdots \\ \mathbf{h}_q^T(L) & \cdots & \mathbf{h}_q^T(L-K) \\ \vdots & \ddots & \vdots \\ \mathbf{0}^T & \cdots & \mathbf{h}_q^T(L) \end{bmatrix}. \quad (24)$$

Consider (20) in the noise-free case and form the $(N-K) \times M(K+1)$ block Hankel data matrix

$$\mathbf{X} := \begin{bmatrix} \mathbf{x}^T(N-1) & \cdots & \mathbf{x}^T(N-1-K) \\ \vdots & \vdots & \vdots \\ \mathbf{x}^T(K) & \cdots & \mathbf{x}^T(0) \end{bmatrix} = \mathbf{S}_b \mathbf{H} \quad (25)$$

where the $(N-K) \times Q(L+K+1)$ modulated input matrix $\mathbf{S}_b$ and the $Q(L+K+1) \times M(K+1)$ channel matrix $\mathbf{H}$ are given by

$$\mathbf{S}_b := \underbrace{\begin{bmatrix} \mathbf{s}_1^T(N-1) & \cdots & \mathbf{s}_Q^T(N-1) \\ \vdots & \vdots & \vdots \\ \mathbf{s}_1^T(K) & \cdots & \mathbf{s}_Q^T(K) \end{bmatrix}}_{\mathbf{S}_1 \quad \cdots \quad \mathbf{S}_Q},$$

$$\mathbf{H} := \begin{bmatrix} \mathbf{H}_1 \\ \vdots \\ \mathbf{H}_Q \end{bmatrix}. \quad (26)$$

In (26), $\{\mathbf{S}_q\}_{q=1}^Q$ are $(N-K) \times (L+K+1)$ Hankel matrices constructed from $\exp(j\omega_q n)s(n)$, $n = -L, \cdots, N-1$. Under the following assumptions it will become possible to estimate the channel matrix $\mathbf{H}$ up to a $Q \times Q$ matrix ambiguity $\mathbf{\Phi}$, which agrees with the fact that for the TI case ($Q = 1$) the ambiguity is a scalar.

- a4) $N - K \geq M(K+1)$, which is easily satisfied by collecting sufficient data;
- a5) $\mathbf{H}$ is at least fat, i.e., the quadruplet (M, L, Q, K) obeys
$$M(K+1) > Q(L+K+1). \quad (27)$$
To satisfy (27), a minimum $M_{\min} = Q+1$ channels are required with a minimum equalizer order $K_{\min} = QL - 1$ (in the TI case, $M_{\min} = 2$ and $K_{\min} = L - 1$ [45], [52]).
- a6) $\mathbf{H}$ is full rank, i.e., rank$(\mathbf{H}) = Q(L+K+1)$ which requires that transfer functions $\{H_q^{(m)}(z), m \in [1, M]\}$ are coprime for every fixed q. This is because, if the family of polynomials $\{H_q^{(m)}(z), m \in [1, M]\}$ have common factors for some q, then, $\mathbf{H}_q$ will lose rank (see e.g., [45]), and hence $\mathbf{H}$ will have linearly dependent rows.
- a7) Bases $\exp(j\omega_q n)$ are sufficiently varying and $s(n)$ is persistently exciting (p.e.) of sufficient order to assure that rank$(\mathbf{S}_b) = Q(L+K+1)$. We stress that $s(n)$ can be either random or deterministic.

To determine $\mathbf{H}$ within the matrix ambiguity $\mathbf{\Phi}$, let us consider (25) and the eigendecomposition

$$\mathbf{X}'\mathbf{X} = \mathbf{H}'\mathbf{S}_b'\mathbf{S}_b\mathbf{H} = [\mathbf{U}_s \mathbf{U}_0] \begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U}_s' \\ \mathbf{U}_0' \end{bmatrix}. \quad (28)$$

Under a7), $\mathcal{R}(\mathbf{U}_s) = \mathcal{R}(\mathbf{H}')$. Since the signal subspace is orthogonal to the noise subspace, if $\mathbf{u} \in \mathcal{R}(\mathbf{U}_0)$, then $\mathbf{Hu} = \mathbf{0}$, and using (26), $\mathbf{H}_q\mathbf{u} = \mathbf{0}$, for $q = 1, \cdots, Q$. Since $\mathbf{H}_q$ is a convolution operator, this can also be written as $\mathcal{T}(\mathbf{u})\mathbf{h}_q = \mathbf{0}$, implying that $\mathbf{h}_q \in \mathcal{N}(\mathcal{T}(\mathbf{u}))$ $q = 1, \cdots, Q$, where $\mathcal{T}(\mathbf{u})$ is a block Toeplitz matrix as defined in (24). Hence, if $\{\tilde{\mathbf{h}}_q\}_{q=1}^Q$ are basis vectors for $\mathcal{N}(\mathcal{T}(\mathbf{u}))$, it follows that there exists a full-rank $Q \times Q$ matrix $\mathbf{\Phi}$ such that

$$[\tilde{\mathbf{h}}_1 \cdots \tilde{\mathbf{h}}_Q]\mathbf{\Phi} = [\mathbf{h}_1 \cdots \mathbf{h}_Q]. \quad (29)$$

Let $\{\tilde{\mathbf{h}}_q\}_{q=1}^Q$ denote the null eigenvectors of $\mathcal{T}(\mathbf{u})$, and $\tilde{\mathbf{H}}_q = \mathcal{T}(\tilde{\mathbf{h}}_q)$ and $\tilde{\mathbf{H}}$ be constructed exactly like $\mathbf{H}_q$ and $\mathbf{H}$ in (24) and (26). If we deconvolve the data with $\tilde{\mathbf{G}} := \tilde{\mathbf{H}}^\dagger$, we obtain: $\mathbf{X}\tilde{\mathbf{G}} = \mathbf{S}_b(\mathbf{\Phi} \otimes \mathbf{I}_{L+K+1})\tilde{\mathbf{H}}\tilde{\mathbf{G}} = \mathbf{S}_b(\mathbf{\Phi} \otimes \mathbf{I}_{L+K+1})$. The latter implies that $\tilde{\mathbf{S}}_b = [\tilde{\mathbf{S}}_1 \cdots \tilde{\mathbf{S}}_Q] := \mathbf{S}_b(\mathbf{\Phi} \otimes \mathbf{I}_{L+K+1})$ is such that

$$\tilde{\mathbf{S}}_j = \sum_{i=1}^{Q} \mathbf{S}_i \phi_{ij}, \qquad j = 1, \cdots, Q \quad (30)$$

where ϕ_{ij} is the (i, j)th entry of Φ. Keeping in mind that $\mathbf{S}_q$ is a matrix formed by the modulated input sequence $s(n)\exp(j\omega_q n)$, we can write (30) as

$$\tilde{\mathbf{s}}(n) := [\tilde{s}_1(n) \cdots \tilde{s}_Q(n)]^T = \mathbf{\Phi}\xi(n)s(n) \quad (31)$$

where $\tilde{s}_q(n)$ is an entry of $\tilde{\mathbf{S}}_q$ and $\xi(n) := [\exp(j\omega_1 n) \cdots \exp(j\omega_Q n)]^T$. Equation (31) can be rearranged to obtain

$$s^{-1}(n)\tilde{\mathbf{s}}(n) - \sum_{q=1}^{Q} \boldsymbol{\phi}_q e^{j\omega_q n} = 0 \tag{32}$$

where $\boldsymbol{\phi}_q$ is the qth column of $\boldsymbol{\Phi}$. Given $\{\tilde{\boldsymbol{s}}(n)\}_{n=0}^{N-1}$, (31), and therefore (32), has a unique solution for $s(n)$ and $\boldsymbol{\Phi}$ (see Appendix II for a proof). Hence, (32) can be cast in a matrix form to obtain $s(n)$ and $\boldsymbol{\Phi}$, which after using (29) yields $\mathbf{H}$.

In summary, to estimate $\mathbf{H}$ and the input, we need to perform a singular value decomposition (SVD) on $\mathbf{X}'\mathbf{X}$ to find the vector $\mathbf{u}$ corresponding to the minimum eigenvalue of $\mathbf{X}'\mathbf{X}$. Then another SVD is performed on $\mathcal{T}(\mathbf{u})$, and the Q vectors corresponding to its minimum Q singular values yield estimates of the ambiguous channels $\{\tilde{\mathbf{h}}_q\}_{q=1}^{Q}$. Upon constructing $\tilde{\mathbf{H}}$ as in (24), we deconvolve the data by computing $\mathbf{X}\tilde{\mathbf{H}}^{\dagger}$ to obtain $\tilde{\mathbf{s}}(n)$. Based on $\tilde{\mathbf{s}}(n)$, the unique solution of (32) can be found by casting it in a matrix form to obtain both $\boldsymbol{\Phi}$ and the input estimates.

This method requires at least one vector in the noise subspace, so $M(K+1) > Q(L+K+1)$ is necessary. An alternative method described in Section V-B allows $\mathbf{H}$ to be square and calculates the columns of its (right) inverse (vectors of equalizer coefficients) directly, using the structure of the input matrix.

D. Order Determination

Up to this point we have assumed that the channel order L and the number of bases Q and K were known. To assert that these blind methods are applicable, one needs to show that it is possible to obtain these quantities from output data. Using the rank properties of the output data matrix $\mathbf{X}$ in (22), it is possible to obtain the channel order L and the number of bases Q [10], [31].

Under a4)–a7), matrix $\mathbf{X}$ in (25) has rank $Q(L+K+1)$. With $\tilde{K}_1 > \tilde{K}_2$ denoting known upper bounds on K, corresponding matrices $\tilde{\mathbf{X}}_1$ and $\tilde{\mathbf{X}}_2$ will have rank$(\tilde{\mathbf{X}}_i) = Q(L+\tilde{K}_i+1)$, $i=1, 2$. It is thus possible to select the orders L and Q using

$$Q = \frac{\text{rank}(\tilde{\mathbf{X}}_1) - \text{rank}(\tilde{\mathbf{X}}_2)}{\tilde{K}_1 - \tilde{K}_2}$$
$$L = \frac{\text{rank}(\tilde{\mathbf{X}}_1)}{Q} - (\tilde{K}_1 + 1).$$

With Q, L available, K is chosen to satisfy (29) for a given $M \geq Q+1$.

At low SNR's, noise will make it difficult to discern small significant singular values of $\tilde{\mathbf{X}}_1$ and $\tilde{\mathbf{X}}_2$ from large insignificant ones. More elaborate tests involving information theoretic criteria, such as the AIC, seem possible but are beyond the scope of this paper.

V. Blind Equalization of TV Channels

In this section we will discuss methods for estimating the input. Having the channel estimates available, maximum-likelihood decoding can be used for that purpose. The high computational complexity of Viterbi's algorithm is even more pronounced for the TV model than the TI case since the number of bases Q, as well as the channel length L, affects the computational complexity. A decision feedback scheme has been proposed in [46] in connection with the exponential basis expansion model. Here we will consider linear options: zero-forcing FIR equalizers requiring enough diversity (at least $M > Q$) in Section V-A and minimum mean square error (MMSE) solutions in Section V-B. Optimally weighted equalizers and adaptive algorithms which pertain to the direct blind equalization method are presented in Sections V-C and V-D.

A. Direct Blind Equalization

This method estimates FIR zero-forcing equalizers that yield perfect estimates in the absence of noise without having to estimate the channel first. Similar to the indirect method of Section IV-C no statistical assumptions on the input are made. Estimation of the direct blind equalizers is less computationally demanding than the indirect method. Also, the linear form of the solution in this section enables updating the equalizer estimates adaptively (see Section V-D).

We seek $M \times 1$ FIR zero-forcing equalizers $\{\mathbf{g}_q^{(d)}(k)\}_{k=0}^{K}$ that satisfy (see also Fig. 6)

$$\sum_{k=0}^{K} \mathbf{x}^T(n-k)\mathbf{g}_q^{(d)}(k) = s(n-d)e^{j\omega_q(n-d)}, \quad q = 1, \cdots, Q \tag{33}$$

where $d \in [0, L+K]$ denotes a delay which is inherently nonidentifiable in blind approaches.

To establish existence and uniqueness of such equalizers, we need $\mathbf{H}$ in (26) to be fat or square so that a $\mathbf{G}$ that satisfies $\mathbf{HG} = \mathbf{I}$ exists. The $(q-1)(L+K+1)+d$th column of $\mathbf{G}$ is $[\mathbf{g}_q^{T(d)}(0) \cdots \mathbf{g}_q^{T(d)}(K)]^T$. For the direct blind equalization method, beyond assumptions a4), a6), and a7), required also for indirect channel estimation, we allow $\mathbf{H}$ to be square so that a5) $M(K+1) = Q(L+K+1)$ [$\mathbf{H}$ in (26) is square] is permissible for the method to work.

In order to find the equalizers $\{\mathbf{g}_q^{(d)}(k)\}_{k=0}^{K}$ we first set $n = N-1, \cdots, K$ in (33) and collect equations to obtain

$$\mathbf{X}\mathbf{g}_q^{(d)} = \mathbf{s}_q^{(d)} = \tilde{\mathbf{B}}_q^{(d)}\mathbf{s}^{(d)} \tag{34}$$

where

$$\mathbf{g}_q^{(d)} := \left[\mathbf{g}_q^{(d)T}(0) \cdots \mathbf{g}_q^{(d)T}(K)\right]^T$$
$$\mathbf{s}_q^{(d)} := \left[e^{j\omega_q(N-1-d)}s(N-1-d) \cdots e^{j\omega_q(K-d)}s(K-d)\right]^T$$
$$\tilde{\mathbf{B}}_q^{(d)} := \text{diag}\left[e^{j\omega_q(N-1-d)} \cdots e^{j\omega_q(K-d)}\right]$$
$$\mathbf{s}^{(d)} := [s(N-1-d) \cdots s(K-d)]^T.$$

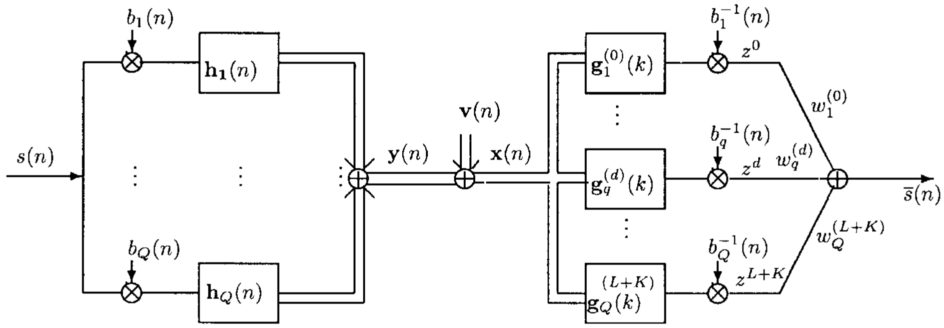

Fig. 6. Vector TV model and FIR vector equalizers.

We use MATLAB's notation $\mathbf{X}(i_1{:}i_2, :)$ to denote a submatrix of $\mathbf{X}$ formed by the i_1 through i_2 rows and all columns of $\mathbf{X}$. So we define

$$\begin{aligned}\mathbf{X}_{0,d} &:= \mathbf{X}(d+1:N-K, :)\\ \mathbf{X}_d &:= \mathbf{X}(1{:}N-K-d, :)\\ \mathbf{B}_q^{(0,d)} &:= \tilde{\mathbf{B}}_q^{(d)}(d+1{:}N-K,\, d+1{:}N-K)\\ \mathbf{B}_q^{(d)} &:= \tilde{\mathbf{B}}_q^{(d)}(1{:}N-K-d,\, 1{:}N-K-d). \end{aligned} \tag{35}$$

Matrix $\mathbf{X}_{0,d}$ is $\mathbf{X}$ without its first d rows, $\mathbf{X}_d$ is $\mathbf{X}$ without its last d rows (likewise for the diagonal matrices, and the modulated input matrices are $\mathbf{S}_q$, to be used in Appendix III).

From (34) and (35) it follows that

$$\begin{aligned}\mathbf{X}_{0,d}\mathbf{g}_{q1}^{(0)} &= \mathbf{B}_{q1}^{(0,d)}\mathbf{s}^{(0)}(d+1{:}N-K)\\ \mathbf{X}_d\mathbf{g}_{q2}^{(d)} &= \mathbf{B}_{q2}^{(d)}\mathbf{s}^{(d)}(1{:}N-K-d).\end{aligned} \tag{36}$$

We note that $\mathbf{s}^{(0)}(d+1{:}N-K) = \mathbf{s}^{(d)}(1{:}N-K-d)$, which allows us to eliminate the input dependence from the equations in (36) and obtain the cross relation

$$\mathbf{B}_{q2}^{(d)}\mathbf{X}_{0,d}\mathbf{g}_{q1}^{(0)} = \mathbf{B}_{q1}^{(0,d)}\mathbf{X}_d\mathbf{g}_{q2}^{(d)}. \tag{37}$$

The pair of equalizers $(\mathbf{g}_{q1}^{(0)}, \mathbf{g}_{q2}^{(d)})$ will be uniquely identifiable (up to a scale) as the eigenvector corresponding to the minimum eigenvalue of $\mathcal{X}_{q1,q2}^{(0,d)}$ in

$$\mathcal{X}_{q1,q2}^{(0,d)}\mathbf{g}_{q1,q2}^{(0,d)} := \left[\mathbf{B}_{q2}^{(d)}\mathbf{X}_{0,d} - \mathbf{B}_{q1}^{(0,d)}\mathbf{X}_d\right]\begin{bmatrix}\mathbf{g}_{q1}^{(0)}\\ \mathbf{g}_{q2}^{(d)}\end{bmatrix} = \mathbf{0} \tag{38}$$

provided that the nullity $\nu(\mathcal{X}_{q1,q2}^{(0,d)}) = 1$. The result is summarized in Theorem 3 (see Appendix III for the proof).

Theorem 3: Under a5), a6), and a7), consider $q_1 = 1$, $q_2 = Q$, $d = L+K$, and $\exp(j\omega_{q_1}n) = \exp(j\omega_{q_1}(n+L+K))$. It then holds that $\nu(\mathcal{X}_{q1,q2}^{(0,d)}) = 1$, and hence (38) has a unique solution. If instead of a5), $M(K+1) > Q(L+K+1)$ holds ($\mathbf{H}$ is fat), then $\nu(\mathcal{X}_{q1,q2}^{(0,d)}) = 2[M(K+1) - Q(L+K+1)] + 1 > 1$, and all vectors in the null space of $\mathcal{X}_{q1,q2}^{(0,d)}$ yield equalizers which, when convolved with the output data, yield perfect input estimates up to a multiple of a known complex exponential sequence in the absence of noise. □

The periodicity requirement on $\exp(j\omega_{q_1}n)$ assumed in Theorem 3 can always be satisfied if $\omega_{q_1} = \omega_1$. This is possible by using the techniques in Section III-B, with which we can easily infer the lowest frequency ω_1. Multiplying both sides of (20) with $\exp(-j\omega_1 n)$ we can "shift" all frequencies by ω_1, so that the first basis function of $\mathbf{x}(n)\exp(-j\omega_1 n)$ will be $\exp(j0n) = 1\,\forall n$, which is periodic with any period.

Requiring $q_1 = 1$, $q_2 = Q$, and $d = L+K$ only enables us to find $\mathbf{g}_{1,Q}^{(0,L+K)}$, but this is not a real concern since, using $\mathbf{g}_{1,Q}^{(0,L+K)}$, equalizers corresponding to other delays and bases (other columns of $\mathbf{G}$) can be found using (37).

Strict inequality in a5) causes every equalizer vector $\mathbf{g}_q^{(d)}$ to lie in an affine space [the set of all vectors $\mathbf{g}$ satisfying $\mathbf{H}\mathbf{g} = \mathbf{e}_{(q-1)(L+K+1)+d+1}$, where $\mathbf{e}_r$ is a unit vector with a 1 in its rth position]. This gives us more freedom in choosing the appropriate equalizer with good noise suppression characteristics. As mentioned in [14] and [15], if the noise is white, the equalizer with the minimum norm will have minimum noise variance at its output.

B. Cylic MMSE Equalizers

Consider the I/O relation in (20). We wish to find, for each n, a vector $\mathbf{g}(n) := [\mathbf{g}^T(n; -K_1)\cdots\mathbf{g}^T(n; K_2)]^T$ so that the following MSE is minimized:

$$E\left|\sum_{k=-K_1}^{K_2}\mathbf{g}'(n;k)\mathbf{x}(n-k) - s(n)\right|^2. \tag{39}$$

The orthogonality principle yields

$$E\left[\sum_{k_1=-K_1}^{K_2}\mathbf{g}'(n;k_1)\mathbf{x}(n-k_1) - s(n)\right]\mathbf{x}'(n-k_2) = 0,\quad k_2 = -K_1,\cdots,K_2. \tag{40}$$

We need to write the set of linear equations in (40) in matrix form and in terms of the estimated channel parameters $\{\mathbf{h}_q(l),\, q = 1,\cdots,Q\}$ and frequencies $\{\omega_q\}_{q=1}^Q$. To this end, we define the following:

$$\begin{aligned}\underbrace{\mathbf{X}(n)}_{(K_1+K_2+1)\times M} &:= [\mathbf{x}(n+K_1)\cdots\mathbf{x}(n-K_2)]^T\\ \underbrace{\overline{\mathbf{H}}_q}_{(L+1)\times M} &:= [\mathbf{h}_q(0)\cdots\mathbf{h}_q(L)]^T\end{aligned} \tag{41}$$

$$\mathbf{B}_q(n) := \mathrm{diag}\Big(e^{j\omega_q(n+K_1)}, \dots, e^{j\omega_q(n-K_2)}\Big) \tag{42}$$

$$\mathbf{S}(n) := \begin{bmatrix} s(n+K_1) & \cdots & s(n+K_1-L) \\ \vdots & \vdots & \vdots \\ s(n-K_2) & \cdots & s(n-K_2-L) \end{bmatrix} \tag{43}$$

Using the definitions (41) and (43), it is easy to verify from (20) that the following equation holds:

$$\mathbf{X}(n) = \sum_{q=1}^{Q} \mathbf{B}_q(n)\mathbf{S}(n)\overline{\mathbf{H}}_q + \mathbf{V}(n) \tag{44}$$

where $\mathbf{V}(n)$ is the corresponding noise matrix. Equation (40) can be rewritten as

$$E\big[\mathrm{vec}(\mathbf{X}^T(n))\mathrm{vec}(\mathbf{X}^T(n))'\big]\mathbf{g}(n) = E\big[\mathrm{vec}(\mathbf{X}^T(n))s^*(n)\big] \tag{45}$$

where $\mathrm{vec}(\mathbf{A})$ denotes the vector formed from concatenating the columns of $\mathbf{A}$. We need to find an expression for $\mathrm{vec}(\mathbf{X}(n))$ using (44), and substitute it into (45) in order to write $\mathbf{g}(n)$ in terms of $\mathbf{B}_q(n)$ and $\mathbf{H}_q$. Taking the transpose of (44) we have $\mathrm{vec}(\mathbf{X}^T(n)) = \mathrm{vec}(\sum_{q=1}^{Q} \overline{\mathbf{H}}_q^T \mathbf{S}^T(n)\mathbf{B}_q^T(n) + \mathbf{V}^{\mathbf{T}}(n)) = \sum_{q=1}^{Q} \mathrm{vec}(\overline{\mathbf{H}}_q^T \mathbf{S}^T(n)\mathbf{B}_q^T(n)) + \mathrm{vec}(\mathbf{V}^T(n))$. Using the identity $\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\mathrm{vec}(\mathbf{B})$, we can conclude the above string of equalities by

$$\mathrm{vec}\big(\mathbf{X}^T(n)\big) = \mathcal{H}_b(n)\mathrm{vec}\big(\mathbf{S}^T(n)\big) + \mathrm{vec}\big(\mathbf{V}^T(n)\big) \tag{46}$$

where $\mathcal{H}_b(n) := \sum_{q=1}^{Q} (\mathbf{B}_q(n) \otimes \overline{\mathbf{H}}_q^T)$. Assuming the noise is white, and using (45) and (46) we have

$$\mathbf{g}(n) = \Big[\mathcal{H}_b(n)E\Big[\mathrm{vec}\big(\mathbf{S}^T(n)\big)\mathrm{vec}\big(\mathbf{S}^T(n)\big)'\Big]\mathcal{H}_b(n)' + \sigma_v^2\mathbf{I}\Big]^{-1}\mathcal{H}_b(n)E\big[\mathrm{vec}\big(\mathbf{S}^T(n)\big)s^*(n)\big]. \tag{47}$$

Next, we define $\mathcal{I} := (1/\sigma_s^2)E[\mathrm{vec}(\mathbf{S}^T(n))\mathrm{vec}(\mathbf{S}^T(n))']$ and using (43) conclude that $\mathcal{I}$ is an $(L+1)(K_1+K_2+1) \times (L+1)(K_1+K_2+1)$ symmetric block Toeplitz matrix with a first block-row $[\mathbf{I}\ \mathbf{I}^{(1)}\ \cdots\ \mathbf{I}^{(L)}\ \cdots\ \mathbf{0}]$, where $\mathbf{I}^{(i)}$ denotes an $(L+1) \times (L+1)$ matrix with ones on the ith subdiagonal, and zero elsewhere. If we also define $\mathbf{e} := (1/\sigma_s^2)E[\mathrm{vec}(\mathbf{S}^T(n))s^*(n)]$, we can express (47) as

$$\mathbf{g}(n) = \left[\mathcal{H}_b(n)\mathcal{I}\mathcal{H}_b(n)' + \frac{\sigma_v^2}{\sigma_s^2}\mathbf{I}\right]^{-1}\mathcal{H}_b(n)\mathbf{e}. \tag{48}$$

One can also obtain a closed form expression for the MMSE after substituting (48) into (39)

$$E|e(n)|^2 = \sigma_s^2\left[1 - \mathbf{e}'\mathcal{H}_b(n)'\left[\mathcal{H}_b(n)\mathcal{I}\mathcal{H}_b(n)' + \frac{\sigma_v^2}{\sigma_s^2}\mathbf{I}\right]^{-1} \cdot \mathcal{H}_b(n)\mathbf{e}\right]. \tag{49}$$

In (47), an inversion takes place for each value of n, unless the frequencies are commensurate, in which case the matrix to be inverted is periodic.

C. Weighted Equalizers

Even when $\mathbf{H}$ is square, the presence of multiple equalizers corresponding to different delays and bases can be used to improve the multiple input estimates in the presence of noise. By aligning, demodulating, and performing weighted combinations of the estimated columns of $\mathbf{S}_b$, one may get better input estimates than using a single equalizer. Let

$$\overline{s}(n) = \sum_{q=1}^{Q}\sum_{d=0}^{L+K} w_q^{(d)} e^{-j\omega_q n}\left[\sum_{k=0}^{K} \mathbf{x}^T(n+d-k)\mathbf{g}_q^{(d)}(k)\right]. \tag{50}$$

We wish to minimize the cost function $J(\mathbf{w}) := (1/N)\lim_{N\to\infty}\sum_{n=0}^{N-1} E|\overline{s}(n) - s(n)|^2$ with respect to the vector of weights $\mathbf{w} := [w_1^{(0)} \cdots w_1^{(L+K)} \cdots w_Q^{(0)} \cdots w_Q^{(L+K)}]^T$, which after using (33) is

$$J(\mathbf{w}) = \frac{1}{N}\lim_{N\to\infty} E\left|\sum_{q=1}^{Q}\sum_{d=0}^{L+K} w_q^{(d)} e^{-j\omega_q n} \cdot \sum_{k=0}^{K} \mathbf{v}^T(n+d-k)\mathbf{g}_d(k)\right|^2. \tag{51}$$

If we constrain the sum of the elements of $\mathbf{w}$ to be one, then $\hat{\mathbf{w}}$ that minimizes (51) is given by (see Appendix IV)

$$\hat{\mathbf{w}} = \frac{1}{\mathbf{1}^T\mathbf{A}^{-1}\mathbf{1}}\mathbf{A}^{-1}\mathbf{1} \tag{52}$$

where $\mathbf{1}$ is a vector of all ones, $\mathbf{A}$ is a block diagonal matrix containing the matrices $\{\mathbf{A}_q\}_{q=1}^{Q}$ in its diagonal, the (i, j) element of which is

$$[\mathbf{A}_q]_{ij} = \sum_{k_1, k_2=0}^{K}\Big[\mathbf{g}_q^{T(i)}(k_1)\mathbf{R}_v(j-i+k_1-k_2) \cdot \mathbf{g}_q^{(j)}(k_2) + \mathbf{g}_q^{T(j)}(k_1) \cdot \mathbf{R}_v(i-j+k_1-k_2)\mathbf{g}_q^{(i)}(k_2)\Big] \tag{53}$$

and $\mathbf{R}_v(\tau) := E[\mathbf{v}(n)\mathbf{v}'(n+\tau)]$.

Notice that the optimum weights require the knowledge of perfect equalizer values which cannot be obtained with noisy data. But with sufficiently high SNR, which enables accurate equalizer estimates, the use of the weights often improves the input estimates, as verified in our simulation examples.

D. Adaptive Equalization

One advantage of the direct method of Section V-B over the indirect method of Section IV-D is the fact that equalizer estimates can be linearly related with the output data, and can be cast into an adaptive framework. The adaptive method proposed to estimate the frequencies in [46] can be combined with what follows to construct an algorithm where both the basis frequencies and the channel parameters can be estimated online.

Equation (38) can be recast in a least squares framework by setting the first coefficient of $\mathbf{g}_{q1,q2}^{(0,L+K)}$ to one and can be rewritten as $\overline{\mathcal{X}}\overline{\mathbf{g}} = \overline{\mathbf{x}} = -\mathbf{x}_1$, where $\overline{\mathcal{X}}$ is $\mathcal{X}_{q1,\,q2}^{(0,\,L+K)}$ without its first column, $\mathbf{x}_1$ is the vector containing the elements of that column, and $\overline{\mathbf{g}}$ is $\mathbf{g}_{q1,\,q2}^{(0,\,L+K)}$ without its first element. It is well known that RLS is a recursive way of computing $\overline{\mathbf{g}}_{LS} = (\overline{\mathcal{X}}'\overline{\mathcal{X}})^{-1}\overline{\mathcal{X}}'\overline{\mathbf{x}}$, which also solves the least squares problem [22]. We use this algorithm to update the vector of equalizer coefficients.

One could also be interested in using the computationally less intensive LMS algorithm at the expense of less accuracy and slower convergence. In the absence of a training sequence (desired input), we consider the elements of $\mathbf{x}_1$ as our desired sequence that we would like $\hat{\mathbf{g}}_t'\hat{\boldsymbol{\eta}}_t$ to estimate. Here, $\hat{\boldsymbol{\eta}}_t$ are the rows of $\overline{\mathcal{X}}$ and $\hat{\mathbf{g}}_t$ is the estimate of the vector of equalizer coefficients at time t. At each iteration, the vector of equalizer coefficients is updated by the relations

$$\hat{\mathbf{g}}_{t+1} = \hat{\mathbf{g}}_t + \mu\epsilon_{t+1}^*\hat{\boldsymbol{\eta}}_t \qquad \epsilon_t = \hat{\mathbf{g}}_t'\hat{\boldsymbol{\eta}}_t - x_{1t} \tag{54}$$

where μ is the step size parameter and x_{1t} denotes tth scalar entry of $\mathbf{x}_1$.

It should be noted that rapid variations of the channel are taken care of by the bases, whereas slower changes in the parameters are tracked by the adaptive algorithm. Since the variation is built in the model, the algorithms can operate on longer data records with less worry about violating the stationarity assumption.

VI. Simulations

In this section we illustrate some of the methods and algorithms that are discussed and compare them. For this purpose we will need the following definitions: the output SNR is defined as $\text{SNR} := \sum_{n=0}^{N-1} \|\mathbf{y}(n)\|^2 / \sum_{n=0}^{N-1} \|\mathbf{v}(n)\|^2$, where $\mathbf{y}(n)$ is the noise-free output data, and the normalized root mean square error (RMSE) between a vector $\mathbf{a}$ and its estimate $\hat{\mathbf{a}}$ is computed as follows:

$$\text{RMSE} = \sqrt{\frac{1}{R}\sum_{r=1}^{R} \frac{\|\hat{\mathbf{a}}_r - \mathbf{a}_r\|^2}{\|\mathbf{a}_r\|^2}} \tag{55}$$

where r stands for realization and R is the number of realizations.

Unless otherwise indicated, $M = 3$ sensors were used with a channel order $L = 3$. The $Q = 2$ bases were chosen as $b_1(n) = 1$ and $b_2(n) = \exp(j(2\pi/50)n)$. All plots except the eye diagrams are an average of 100 Monte Carlo runs unless otherwise indicated.

In Fig. 7, we illustrate how the blind algorithm that is developed in [15] for TI channels compares with the one proposed in Section V-A, when the data comes from a rapidly fading TV channel. We see that the TI algorithm is not capable of equalizing the $N = 150$ symbols coming from a 16-QAM constellation even with a high SNR of 45 dB. This motivating example demonstrates the inadequacy of TI equalization algorithms when applied to TV channels.

Fig. 8 illustrates the frequency estimation of Section III-B using $N = 1000$ data at an SNR = 10 dB. Since there was only one nonzero cycle $2\pi/50$, $m_{11}(n;\,0)$ provided enough

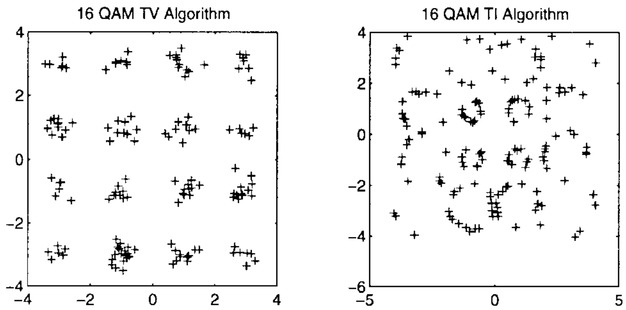

Fig. 7. TI and TV algorithms on TV data.

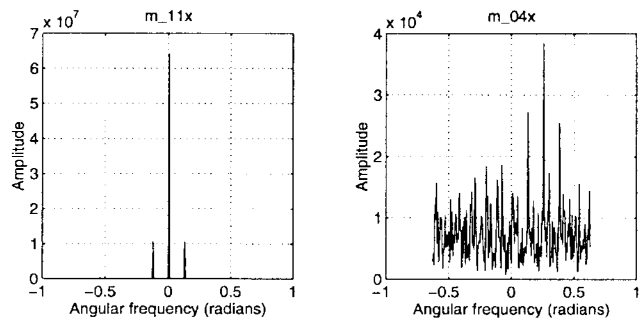

Fig. 8. Estimation of basis frequencies.

information to estimate it. The plot on the left shows that the Fourier Series of $m_{11_x}(n;\,0)$ has two peaks: one at $\omega = 0$ (due to the stationary noise) and the other $\omega = 2\pi/50$. The plot on the right illustrates the estimation of the same cycle with fourth-order cyclic moment at lag $\tau = 0$ computed as the FFT of $x^4(n)$. We observe peaks at multiples of $2\pi/50$. The reduced variance of second order statistics relative to fourth-order statistics is also apparent.

In Fig. 9, the five least significant singular values of the matrix $\mathbf{X}$ are plotted for $K = 8$ (left) and $K = 9$ (right) and for SNR's of 50 and 25 dB. Only $N = 150$ samples were used. The number of least significant singular values (zero singular values in the case of no noise) determines the rank of the noise free output data matrix, which, as shown in Section IV-D, enables the estimation of L, Q, and K. With an SNR $= 50$ dB, the insignificant singular values are still discernible. As the SNR's get lower (to 25 dB), it becomes more difficult to tell how many zero singular values there are since the noise not only increases them but also perturbs their relative values. The standard deviation of the singular values is also plotted around the mean which was estimated from 100 realizations.

Fig. 10 illustrates the two sensor approach of Section IV-B, where the channel coefficients are estimated from the cyclic correlations of the output. The RMSE between the true channel coefficients corresponding to $\{h_1^{(m)}(l)\ m = 1,\,2\}$ is plotted versus SNR (200 Monte Carlo runs, $N = 1000$) and the number of data (500 Monte Carlo runs, SNR $=$ 20 dB). We see that (unlike the deterministic methods) the channel estimates are consistent and improve significantly with the number of data. In addition, the effect of noise is minimal due to the use of nonzero cycles in the cyclic correlations.

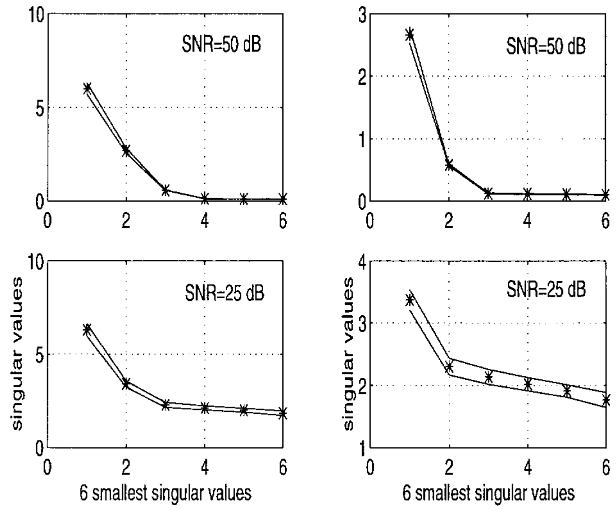

Fig. 9. Order determination.

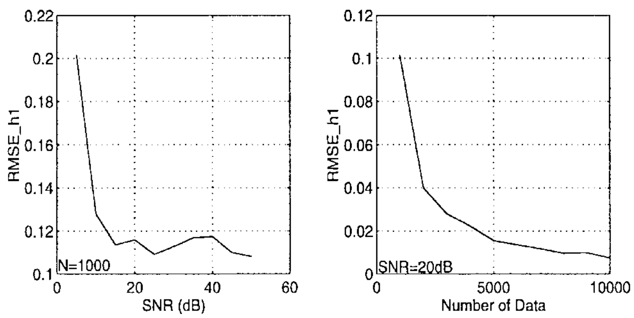

Fig. 10. Two sensor (statistical) approach.

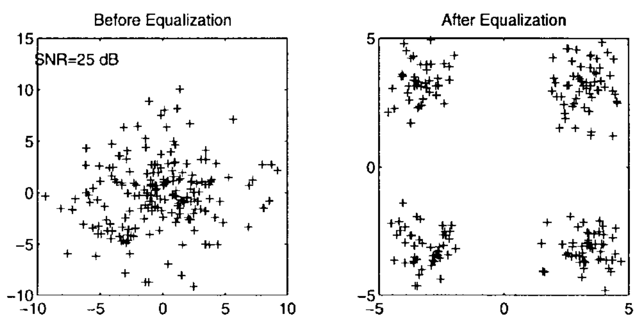

Fig. 11. Before and after equalization.

Fig. 11 illustrates MMSE equalization for the estimated channel where the channel estimation was done with the two-sensor approach. We show the eye diagram for the symbol estimates at an SNR $= 25$ dB. The unequalized channel output is shown on the left plot; the right plot is obtained by using the channel estimates obtained with $N = 5000$ data points and then using the cyclic MMSE equalizer of length 15 [$K_1 = 7$, $K_2 = 7$ in (39)].

In Fig. 12 the deterministic methods are compared. The direct method is implemented in two different ways. The first one, referred to as "direct," uses what is suggested right before (38). The "min-norm" approach substitutes the "direct" $\mathbf{g}_1^{(0)}$ estimate in (37) and solves (36) with respect to $\mathbf{g}_Q^{(d)}$ constraining it to have minimum norm. MSE of

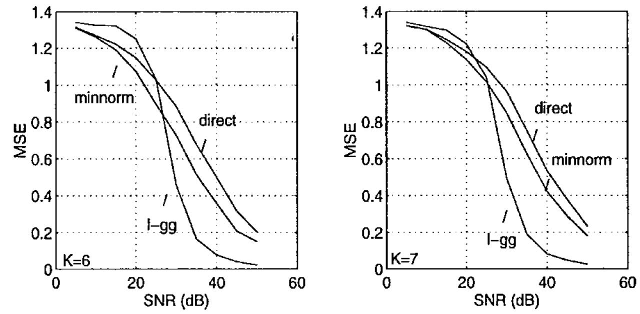

Fig. 12. Direct, indirect, and minimum norm methods versus SNR.

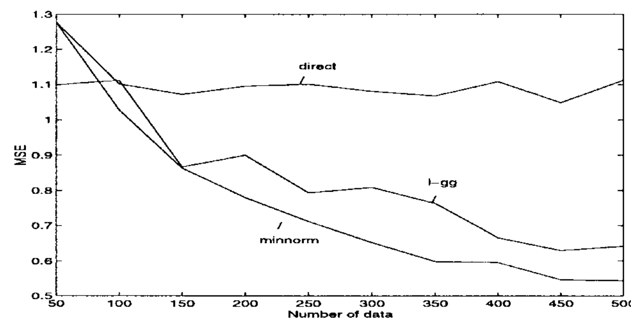

Fig. 13. Performance versus number of data.

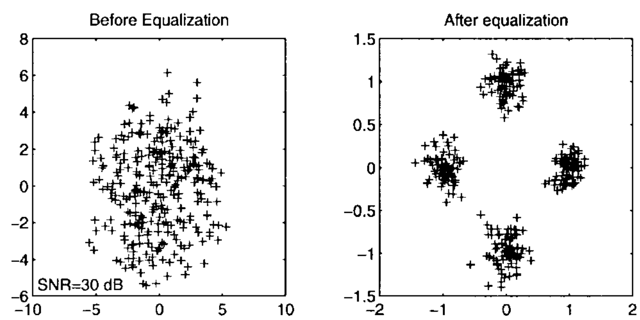

Fig. 14. RLS algorithm, approximate initialization.

the input estimates are plotted for two different equalizer lengths $K = 6, 7$. The number of data used for these estimates was only $N = 100$. Equalizers with minimum norm consistently outperform those obtained using (38). The direct and minimum norm methods perform better than the method in [31] (l-gg) for low (<25 dB) SNR's.

To see how much the deterministic methods improve with increased data length, Fig. 13 compares the minimum norm, direct, and indirect methods. It is seen that the minimum norm method benefits from the increase of the data length more consistently than the other two methods. An SNR of 25 dB was used.

Figs. 14–16 illustrate the performance of the adaptive algorithms proposed in Section V-D. Here $M = 8$, and $K = 0$ and $L = 3$ were chosen. Fig. 14 shows the eye diagrams for the output of an equalizer obtained with the RLS algorithm. Fig. 15 illustrates the performance of the RLS algorithm by plotting the error of the equalizer estimates and also the error in the input estimates. Fig. 16

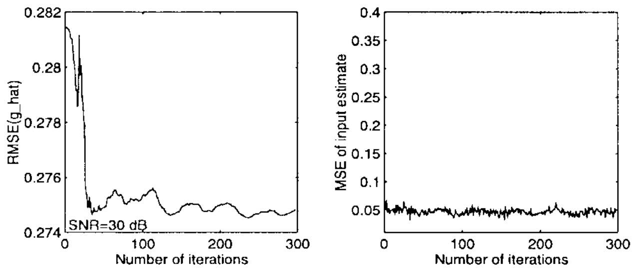

Fig. 15. RLS with the number of iterations.

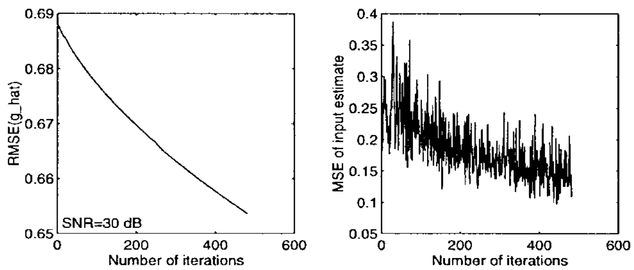

Fig. 16. Performance of the LMS algorithm.

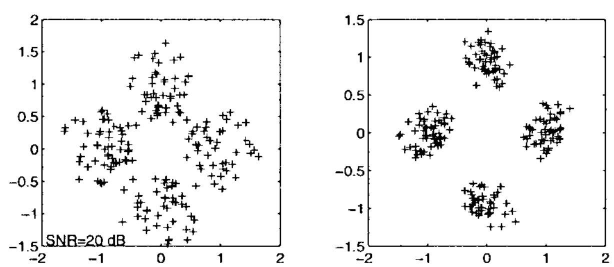

Fig. 17. Zero-delay versus average equalizers.

is the same as Fig. 15 except for the LMS algorithm. The RLS was initialized by $\mathbf{0}$, whereas the LMS was initialized with the batch estimate obtained with the minimum number of symbols N required by a4).

In Fig. 17 effects of weighting of different equalizers on the input estimates are demonstrated. Here the equalizers weighted by the inverse of their norms (right) yielded better estimates than $\mathbf{g}_1^{(0)}$, the zero-delay equalizer (left).

These preliminary simulations illustrate the difference between the "statistical" and the "deterministic" approaches for the TV model we have justified and adopted. While the former relies on cyclic correlations and is effected minimally by the presence of noise, it needs relatively long data records for accurate estimates. The zero-forcing FIR solutions, on the other hand, yield good estimates at high SNR's with short data records, but their noise tolerance is rather small.

VII. Conclusions and Future Directions

Finitely parameterized basis expansions turn SISO TV systems into multivariate TI systems with inputs formed by modulating a single input with the bases. Fourier bases are well motivated for modeling rapidly fading mobile communication channels when multipath propagation caused by a few dominant reflectors gives rise to (Doppler-induced) linearly varying path delays. Doppler frequencies can be estimated blindly using cyclic statistics, and channel orders can be determined from rank properties of a received data matrix.

Structured variations described by bases offer TV channel diversity, which renders blind identification of TV models easier than that of TI models. When channel (or Doppler) diversity is complemented by temporal or spatial diversity (available with oversampling or multiple antennas), blind estimators of TV channels along with direct equalizers become available even with minimal (persistence-of-excitation) assumptions about the input and the bases. The equalizers are TI, multivariate, zero-forcing (in the absence of noise), and lend themselves to optimally weighted and adaptive algorithms. The latter provide fine tuning for possible model mismatch of the bases, which capture the nominal part of the rapidly fading channel. Exploitation of the input's whiteness reduces the amount of spatio/temporal diversity (only two sensors) needed to identify blindly TV channels and mitigate their effects using MMSE equalizers.

The blind channel estimation and identification methods presented in this paper relied on second-order output information only. In [47], blind higher order statistical methods have been developed which rely on the independence of the input but are capable of identifying TV channels using single sensor data only. Following the start-up mode, blind methods switch on to a decision-directed mode. Decision feedback equalizers for the TV basis expansion model have been reported in [46] along with adaptive methods for on-line estimation of the basis frequencies.

A number of interesting directions open up for future research: 1) performance analysis of the channel estimators, especially when model perturbations due to synchronization effects and Doppler frequency drifts are present; 2) theoretical evaluation in terms of error probability for the zero-forcing equalizers and experimental comparisons with the MSE equalizers; 3) extensions of blind methods to TV pole-zero channel models; 4) exploitation of input redundancy in the form of short training sequences (semi-blind extensions), modulation, codes, or filterbanks in order to identify TV-basis expansion models without oversampling or deployment of multiple antennas. Such input-diversity techniques have gained popularity recently for blind identification of TI channels (see [4], [11], [19], [30], [41], and [48], and references therein); 5) diversity techniques for blind identification of random coefficient models and performance comparisons with the basis expansion models using real data.

Appendix I
Proof of Theorem 1

We will use the notation $\overline{T}(\mathbf{u})$ to denote a convolution matrix with Toeplitz structure associated with the vector $\mathbf{u}$, whose first column is $[\mathbf{u}^T \cdots 0]^T$, and first row is $[u(0) \ \cdots \ 0]_{1\times(L+1)}$, where $u(0)$ is the first element of $\mathbf{u}$. The dimensions depend on the size of the vector that $\overline{T}(\mathbf{u})$ is multiplying and will be clear from the context.

a) Let ω_{q_2} and ω_{q_1} be a pair such that $S_{xx}^{(m_1,m_2)}(\omega_{q_2} - \omega_{q_1}; z) = \sigma_s^2 H_{q_2}^{*(m_2)}(z^* e^{-j\omega_{q_2}}) H_{q_1}^{(m_1)}(e^{-j\omega_{q_2}}/z)$ in (23). Then, since $H_{q_1}^{(1)}(z)$ and $H_{q_1}^{(2)}(z)$ are coprime, the following holds:

$$\frac{S_{xx}^{(1,1)}(\omega_{q_2} - \omega_{q_1}; z)}{S_{xx}^{(2,1)}(\omega_{q_2} - \omega_{q_1}; z)} = \frac{H_{q_1}^{(1)}(e^{-j\omega_{q_2}}/z)}{H_{q_1}^{(2)}(e^{-j\omega_{q_2}}/z)}. \tag{56}$$

Define $\mathbf{c}_{xx}^{(m_1,m_2)}(\alpha) := [C_{xx}^{(m_1,m_2)}(\alpha; -L) \cdots C^{(m_1,m_2)}(\alpha; L)]^T$, the $(3L+1) \times (L+1)$ matrix $\mathbf{C}_{xx}^{(m,1)} := \overline{\mathcal{T}}(\mathbf{c}_{xx}^{(m,1)}(\omega_{q_2} - \omega_{q_1}))$, and $\mathbf{h}_{q_1}^{(m)} := [h_{q_1}^{(m)}(L) e^{j\omega_{q_2} L} \cdots h_{q_1}^{(m)}(0) e^{j\omega_{q_2} 0}]^T, m = 1, 2$. Cross-multiplying in (56), taking the inverse z-transform of both sides, and casting the resulting convolutions in matrix form we obtain

$$\begin{bmatrix} \mathbf{C}_{xx}^{(1,1)} & -\mathbf{C}_{xx}^{(2,1)} \end{bmatrix} \begin{bmatrix} \mathbf{h}_{q_1}^{(2)} \\ \mathbf{h}_{q_1}^{(1)} \end{bmatrix} = \mathbf{0}. \tag{57}$$

The solution in (57) is unique up to a scale, since if $\{\tilde{H}_q^{(m)}(z)\}$ channels also satisfy (56), then

$$\frac{H_{q_1}^{(1)}(e^{-j\omega_{q_2}}/z)}{H_{q_1}^{(2)}(e^{-j\omega_{q_2}}/z)} = \frac{\tilde{H}_{q_1}^{(1)}(e^{-j\omega_{q_2}}/z)}{\tilde{H}_{q_1}^{(2)}(e^{-j\omega_{q_2}}/z)}$$

and the numerators and denominators of both sides must be equal. This is easily shown by factoring the numerator and the denominator polynomials into their L factors. We showed how to estimate $\{h_{q_1}^{(m)}(l), m = 1, 2\}$. Subchannels corresponding to q_2 can be estimated with a similar procedure.

b) We showed in a) how to estimate $\{h_Q^{(m)}(l), h_1^{(m)}(l), m = 1, 2\}$ using $S_{xx}^{(m_1,m_2)}(\omega_Q - \omega_1; z)$, since only the pair ω_Q and ω_1 can give rise to the difference $\omega_Q - \omega_1$. Consider now $S_{xx}^{(m_1,m_2)}(\omega_Q - \omega_2; z)$. If ω_Q, ω_2 is the only pair of frequencies that has the difference $\omega_Q - \omega_2$, then $S_{xx}^{(m_1,m_2)}(\omega_Q - \omega_2; z)$ is the product of two polynomials, one of which is known. This will enable estimating $\{h_2^{(m)}(l), m = 1, 2\}$. If, on the other hand, $\omega_Q - \omega_2 = \omega_{Q-1} - \omega_1$ (the only other pair that could possibly give rise to this difference), then we have the sum of two products of polynomials where two of the four polynomials are known

$$\begin{aligned} &S_{xx}^{(m_1,m_2)}(\omega_Q - \omega_2; z) \\ &= \sigma_s^2 \Big[H_Q^{*(m_2)}(z^* e^{-j\omega_Q}) H_2^{(m_1)}(e^{-j\omega_Q}/z) \\ &\quad + H_{Q-1}^{*(m_2)}(z^* e^{-j\omega_{Q-1}}) H_1^{(m_1)}(e^{-j\omega_{Q-1}}/z) \Big]. \end{aligned} \tag{58}$$

After taking the inverse z-transform of both sides, (58) can be cast in matrix form,

$$\underbrace{\begin{bmatrix} \mathbf{H}_Q^{*(m_2)} & -\mathbf{H}_1^{(m_1)} \end{bmatrix}}_{\mathcal{H}_{Q,1}^{(m_1,m_2)}} \underbrace{\begin{bmatrix} \mathbf{h}_2^{(m_1)} \\ \overline{\mathbf{h}}_{Q-1}^{*(m_2)} \end{bmatrix}}_{\mathbf{h}_{2,Q-1}^{(m_1,m_2)}} = \mathbf{c}_{m_1,m_2} \tag{59}$$

where

$$\begin{aligned} \mathbf{h}_2^{(m)} &:= \left[h_2^{(m)}(L) e^{j\omega_Q L} \cdots h_2^{(m)}(0) e^{j\omega_Q 0} \right]^T \\ \overline{\mathbf{h}}_{Q-1}^{*(m)} &:= \left[h_{Q-1}^{(m)}(0) e^{j\omega_{Q-1} 0} \cdots h_{Q-1}^{(m)}(L) e^{j\omega_{Q-1} L} \right]^{*T} \\ \mathbf{H}_Q^{*(m)} &:= \overline{\mathcal{T}} \left(\left[h_Q^{(m)}(0) e^{j\omega_Q 0} \cdots h_Q^{(m)}(L) e^{j\omega_Q L} \right]^{*T} \right) \\ \mathbf{H}_1^{(m)} &:= \overline{\mathcal{T}} \left(\left[h_1^{(m)}(L) e^{j\omega_{Q-1} L} \cdots h_1^{(m)}(0) e^{j\omega_{Q-1} 0} \right]^T \right) \end{aligned}$$

and vector $\mathbf{c}_{m_1,m_2}$ contains the inverse z-transform of $S_{xx}^{(m_1,m_2)}(\omega_Q - \omega_2; z)$. If $H_Q^{*(m_2)}(z^* e^{-j\omega_Q})$ and $H_1^{(m_1)}(e^{-j\omega_Q}/z)$ are coprime, the $(2L+1) \times (2L+2)$ matrix $\mathcal{H}_{Q,1}^{(m_1,m_2)}$ in (59) has full row rank, which will enable us to determine $\mathbf{h}_{2,Q-1}^{(m_1,m_2)}$ up to a scale ambiguity.

After estimating $\{h_{Q-1}^{(m)}(l), h_2^{(m)}(l), m = 1, 2\}$, we can repeat the same procedure with $\{h_{Q-2}^{(m)}(l), h_3^{(m)}(l), m = 1, 2\}$. This time, $S_{xx}^{(m,k)}(\omega_Q - \omega_3; z)$ might contain a sum of three products, but if this is the case one of the products has to involve $\{h_{Q-1}^{(m)}(l), h_2^{(m)}(l), m = 1, 2\}$ which has been estimated. Proceeding in this fashion, all subchannels can be estimated provided that $H_Q^{*(m_2)}(z^* e^{-j\omega_Q})$ and $H_1^{(m_1)}(e^{-j\omega_{q_1}}/z)$ are coprime for $m_1, m_2 = 1, 2$, whenever there exists a q_2 with $\omega_Q - \omega_{q_2} = \omega_{q_1} - \omega_1$. $\square$

APPENDIX II
UNIQUENESS PROOF OF (32)

As mentioned in Section V-B, without loss of generality we will assume $0 = \omega_1 < \cdots < \omega_Q$. Suppose now, in addition to $s(n)$ and $\mathbf{\Phi}$, $\hat{s}(n)$ and $\hat{\mathbf{\Phi}}$ also satisfy (31). Relating them, we have

$$\underbrace{\hat{\mathbf{\Phi}}^{-1}\mathbf{\Phi}}_{:=\mathbf{F}} \boldsymbol{\xi}(n) = \boldsymbol{\xi}(n) \underbrace{\frac{\hat{s}(n)}{s(n)}}_{:=a(n)} \tag{60}$$

provided that $s(n)$ is nonzero. Equating the first element of both sides in (60), and likewise the last elements, we obtain

$$a(n) = \sum_{q=1}^{Q} F_{1q} e^{j\omega_q n} = \sum_{q=1}^{Q} F_{Qq} e^{j(\omega_q - \omega_Q)n} \, \forall n \tag{61}$$

where F_{ij} denotes the (i, j) element of matrix $\mathbf{F}$ defined in (60). Using the last equality in (61) we can relate the first and last columns of $\mathbf{F}$ and write it for $n = 1, \cdots, N-1$ to obtain the matrix equation (62) as shown at the bottom of the next page. Equation (62) has a unique solution (up to a scale) with $\beta := F_{11} = F_{QQ}$ and $F_{12} = \cdots = F_{Q-1Q} = 0$ due to the Vandermonde structure of $\mathbf{E}$ in (62) and the fact that its first and last columns are identical. This means the first and the last elements of (60) are equal to β and are independent of n. Thus, (60) implies that $a(n) = \beta \, \forall n$, and $\mathbf{F} = \beta\mathbf{I}$; hence, $\mathbf{\Phi} = \beta\hat{\mathbf{\Phi}}$. $\square$

Appendix III
Proof of Theorem 2

Assume first that $\mathbf{H}$ is square. The nullity of

$$\mathcal{X}_{q1,\,q2}^{(0,\,d)} = \underbrace{\left[\mathbf{B}_{q2}^{(d)}[\mathbf{S}_1^{(0,\,d)} \cdots \mathbf{S}_Q^{(0,\,d)}] - \mathbf{B}_{q1}^{(0,\,d)}[\mathbf{S}_1^{(d)} \cdots \mathbf{S}_Q^{(d)}]\right]}_{\mathbf{S}_{q_1,\,q_2}^{(0,\,d)}} \cdot \begin{bmatrix} \mathbf{H} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} \end{bmatrix} \tag{63}$$

is one if and only if the nullity of $\mathbf{S}_{q_1,\,q_2}^{(0,\,d)}$ is one because $\mathbf{H}$ is full rank and square. Because matrix $\mathbf{S}_b$ in (26) is full rank, $\mathbf{S}_{q_1,\,q_2}^{(0,\,d)}$ might lose rank only if columns of one portion are linearly dependent with columns of the other portion. Note that matrices $[\mathbf{S}_1^{(0,\,d)} \cdots \mathbf{S}_Q^{(0,\,d)}]$ and $[\mathbf{S}_1^{(d)} \cdots \mathbf{S}_Q^{(d)}]$ share $Q(L+K-d+1)$ columns. Careful examination of $\mathbf{S}_{q_1,\,q_2}^{(0,\,d)}$ in (63) will reveal that they involve the products $\mathbf{B}_{q2}^{(d)}\mathbf{S}_{q_1}^{(0,\,d)}$ and $-\mathbf{B}_{q1}^{(0,\,d)}\mathbf{S}_{q_2}^{(d)}$, which have $L+K-d+1$ columns in common. So $d = L+K$ is a necessary condition for $\nu(\mathcal{X}_{q1,\,q2}^{(0,\,d)}) = 1$. But other products such as $\mathbf{B}_{q2}^{(L+K)}\mathbf{S}_{q_3}^{(0,\,L+K)}$ and $-\mathbf{B}_{q1}^{(0,\,L+K)}\mathbf{S}_{q_4}^{(L+K)}$ could also have common columns. In fact, if $\exp(j\omega_{q_1}n)$ is periodic with period $L+K$, the first column of $\mathbf{B}_{q2}^{(L+K)}\mathbf{S}_{q_3}^{(0,\,L+K)}$ is identical with the last column of $\mathbf{B}_{q1}^{(0,\,L+K)}\mathbf{S}_{q_4}^{(L+K)}$ whenever $\omega_{q_2}+\omega_{q_3} = \omega_{q_1}+\omega_{q_4}$. Choosing $\omega_{q_1} = \omega_1$ and $\omega_{q_2} = \omega_Q$ will ensure $\omega_{q_3} = \omega_1$, $\omega_{q_4} = \omega_Q$, and along with $d = L+K$, will guarantee that $\nu(\mathbf{S}_{q_1,\,q_2}^{(0,\,d)}) = 1$, therefore $\nu(\mathcal{X}_{q1,\,q2}^{(0,\,d)}) = 1$.

If $M(K+1) > Q(L+K+1)$ ($\mathbf{H}$ is fat), then relying on a6) and using the Sylvester inequality, we get $\text{rank}(\mathcal{X}_{q1,\,q2}^{(0,\,L+K)}) = 2Q(L+K+1)-1$, therefore, $\nu(\mathcal{X}_{q1,\,q2}^{(0,\,L+K)}) = 2M(K+1) - \text{rank}(\mathcal{X}_{q1,\,q2}^{(0,\,L+K)}) = 2[M(K+1)-Q(L+K+1)]+1 > 1$, where the inequality follows from $M(K+1) > Q(L+K+1)$. $\square$

Appendix IV
Derivation of Optimum Weights

We wish to minimize (51) subject to the constraint $\sum_{q=1}^{Q}\sum_{i=0}^{L+K} w_q^{(i)} = 1$. We will use the method of Lagrange multipliers to solve this constrained optimization problem (e.g., [22, pp. 557–560]). The constant σ_v^2 has no affect on the relative weights $w_q^{(i)}$, since the scaling of $\mathbf{w}$ is absorbed into the Lagrange multiplier. By writing (51) as a term times its conjugate and moving the expected value inside we can express it as

$$\sum_{q=1}^{Q}\sum_{i,\,j=0}^{L+K} w_q^{*(i)}w_q^{(j)} \cdot \left[\sum_{k_1,\,k_2=0}^{K} \mathbf{g}_q^{T(i)}(k_1)\mathbf{R}_v(j-i+k_1-k_2)\mathbf{g}_q^{(j)}(k_2)\right] \tag{64}$$

where we used $(1/N)\sum_{n=0}^{N-1}\exp(j(\omega_{q_1}-\omega_{q_2})n) = \delta(\omega_{q_1}-\omega_{q_2})$ and $\mathbf{R}_v(\tau) := E[\mathbf{v}(n)\mathbf{v}'(n+\tau)]$. The objective function to be minimized now becomes

$$\mathcal{J}(\mathbf{w}) = \mathbf{w}^T\mathbf{B}\mathbf{w} - \gamma\mathbf{w}^T\mathbf{1} \tag{65}$$

where γ is the Lagrange multiplier that will determine the scaling of $\mathbf{w}$ to fit the constraint, $\mathbf{B}$ is a block diagonal matrix containing the matrices $\{\mathbf{B}_q\}_{q=1}^{Q}$ in its diagonal, and the $[\mathbf{B}_q]_{ij}$ is given by the sum inside the brackets in (64). Since $\mathbf{B}$ is not Hermitian-symmetric ($\mathbf{B}' \neq \mathbf{B}$), we write

$$\mathbf{B} = \mathbf{A} + \mathbf{D} = \tfrac{1}{2}(\mathbf{B}'+\mathbf{B}) + \tfrac{1}{2}(\mathbf{B}'-\mathbf{B})$$

where $\mathbf{A}$ is Hermitian symmetric and $\mathbf{D}$ is Hermittian anti-symmetric. Because (51), and therefore (64), is real, and due to the Hermitian antisymmetry of $\mathbf{D}$, $\mathbf{w}'\mathbf{D}\mathbf{w}$ is purely imaginary; thus it follows that $\mathbf{w}'\mathbf{D}\mathbf{w} = 0$. Retaining the Hermitian part of $\mathbf{B}$ amounts to an extra term in (53) to guarantee $[\mathbf{A}]_{ij} = [\mathbf{A}]_{ji}^*$. Now we use the standard result $\nabla_{\mathbf{w}}(\mathbf{w}'\mathbf{A}\mathbf{w}) = 2\mathbf{A}\mathbf{w}$ for a Hermitian matrix $\mathbf{A}$ [22] to take the gradient of the objective function in (65) and equate it to zero

$$\nabla_{\mathbf{w}}\mathcal{J}(\mathbf{w}) = 2\mathbf{A}\mathbf{w} - \gamma\mathbf{1} = \mathbf{0}. \tag{66}$$

Equation (66) enables us to solve for $\mathbf{w}$ which, after imposing the constraint, leads to (52) and (53). $\square$

Acknowledgment

The authors wish to thank Prof. H. Liu, former graduate students, and especially Prof. M. K. Tsatsanis of Stevens Institute of Technology for his collaboration and discussions which initiated and formulated a number of topics in this paper.

References

[1] G. D'Aria, F. Muratore, and V. Palestini, "Simulation and performance of the pan-European land mobile radio system," *IEEE Trans. Veh. Technol.*, vol. 41, pp. 177–189, 1992.

$$\underbrace{\begin{bmatrix} 1 & e^{j\omega_2} & \cdots & e^{j\omega_Q} & e^{-j\omega_Q} & e^{j(\omega_2-\omega_Q)} & \cdots & 1 \\ 1 & e^{j2\omega_2} & \cdots & e^{j2\omega_Q} & e^{-j2\omega_Q} & e^{j2(\omega_2-\omega_Q)} & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & e^{jN\omega_2} & \cdots & e^{jN\omega_Q} & e^{-jN\omega_Q} & e^{jN(\omega_2-\omega_Q)} & \cdots & 1 \end{bmatrix}}_{:=\mathbf{E}} \begin{bmatrix} F_{11} \\ \vdots \\ F_{1Q} \\ -F_{Q1} \\ \vdots \\ -F_{QQ} \end{bmatrix} = 0 \tag{62}$$

[2] G. D'Aria, R. Piermarini, and V. Zingarelli, "Fast adaptive equalizers for narrow-band TDMA mobile radio," *IEEE Trans. Veh. Technol.,* vol. 40, pp. 392–404, May 1991.
[3] C. Carlemalm and A. Logothetis, "On detection of double talk and changes in the echo path using a Markov modulated channel model," in *Proc. Intl. Conf. ASSP*, Munich, Germany, Apr. 20–24, 1997, vol. V, pp. 3869–3872.
[4] E. De Carvalho and D. T. M. Slock, "Cramer–Rao bounds for semi-blind and training sequence based channel estimation," in *Proc. 1st IEEE Signal Processing Workshop Wireless Communications,* Paris, France, Apr. 20–24, 1997, pp. 129–132.
[5] 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.
[6] L. Davis, I. Collins, and R. J. Evans, "Identification of time-varying linear channels," in *Proc. Intl. Conf. ASSP,* Munich, Germany, Apr. 20–24, 1997, vol. V, pp. 3921–3924.
[7] *EIA/TIA Interim Standard IS-54, Cellular System Dual-Mode Mobile Station-Base Station Compatibility Standard*, Electronic Industries Association, May 1990.
[8] A. W. Fuxjaeger and R. A. Iltis, "Adaptive parameter estimation using parallel Kalman filtering for spread spectrum code and Doppler tracking," *IEEE Trans. Commun.,* vol. 42, pp. 2227–2230, June 1994.
[9] ——, "Acquisition of timing and Doppler shift in a direct sequence spread spectrum system," *IEEE Trans. Commun.,* vol. 42, pp. 2870–2879, Oct. 1994.
[10] G. B. Giannakis, "Blind equalization of time varying channels: A deterministic multichannel approach," in *Proc. 8th Signal Processing Workshop Statistical Signal Array Processing,* Corfu, Greece, June 24–26, 1996, pp. 180–183.
[11] ——, "Filterbanks for blind channel identification and equalization," *IEEE Signal Processing Lett.,* vol. 4, pp. 184–187, June 1997.
[12] ——, "Cyclostationary signal analysis," *Digital Signal Processing Handbook,* V. K. Madisetti and D. Williams, Eds. Boca Raton, FL: CRC Press, 1998.
[13] G. B. Giannakis and J. M. Mendel, "Identification of nonminimum phase systems using higher-order statistics," *IEEE Trans. Acoustics Speech Signal Processing,* vol. 37, pp. 360–377, Mar. 1989.
[14] G. B. Giannakis and S. Halford, "Blind fractionally-spaced equalization of noisy FIR channels: Direct and adaptive solutions," *IEEE Trans. Signal Processing,* vol. 45, pp. 2277–2292, Sept. 1997.
[15] G. B. Giannakis and C. Tepedelenlioğlu, "Direct blind equalizers of multiple FIR channels: A deterministic approach," submitted for publication.
[16] G. B. Giannakis, C. Tepedelenlioğlu, and H. Liu, "Adaptive blind equalization of time-varying channels," in *Proc. Intl. Conf. ASSP,* Munich, Germany, Apr. 20–24, 1997, vol. V, pp. 4033–4036.
[17] F. Gini and G. B. Giannakis, "Frequency offset and symbol timing recovery in flat fading channels: A unifying cyclostationary approach," submitted for publication.
[18] D. N. Godard, "Self-recovering equalization and carrier-tracking in two-dimensional data communication systems," *IEEE Trans. Commun.,* vol. 28, pp. 1867–1875, 1980.
[19] A. Gorokhov and P. Loubaton, "Semi-blind second-order identification of convolutive channels," in *Proc. Intl. Conf. ASSP*, Munich, Germany, Apr. 20–24, 1997, vol. V, p. 3905.
[20] L. Greenstein and B. Czekaj, "Modeling multipath fading responses using multitone probing signals and polynomial approximation," *Bell Syst. Tech. J.,* vol. 60, pp. 193–214, 1981.
[21] Y. Grenier, "Time-dependent ARMA modeling of nonstationary signals," *IEEE Trans. Audio, Speech, Signal Processing,* vol. 31, pp. 899–911, Aug. 1983.
[22] S. Haykin, *Adaptive Filter Theory.* Englewood Cliffs, NJ: Prentice-Hall, Inc., 1996.
[23] P. Hoeher, "A statistical discrete-time model for the WSSUS multipath channel," *IEEE Trans. Veh. Technol.,* vol. 41, pp. 461–468, Apr. 1992.
[24] R. A. Iltis, "Joint estimation of PN code delay and multipath using extended Kalman filter," *IEEE Trans. Commun.,* vol. 38, pp. 1677–1685, Oct. 1990.
[25] R. A. Iltis and A. W. Fuxjaeger, "A digital DS spread-spectrum receiver with joint channel and Doppler shift estimation," *IEEE Trans. Commun.,* vol. 39, pp. 1255–1267, Aug. 1991.
[26] R. A. Iltis, J. J. Shynk, and K. Giridhar, "Bayesian algorithms for blind equalization using parallel adaptive filters," *IEEE Trans. Commun.,* vol. 42, pp. 1017–1032, Feb.–Apr. 1994.
[27] W. C. Jakes, *Microwave Mobile Communications.* New York: IEEE Press, 1994.
[28] M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Simulation of Communication Systems.* New York: Plenum Press, 1992.
[29] L. A. Liporace, "Linear estimation of nonstationary signals," *J. Acoustic Soc. Amer.,* vol. 58, pp. 1288–1295, 1975.
[30] A. Chevreuil and P. Loubaton, "Blind second-order identification of FIR channels: Forced cyclostationarity and structured subspace method," in *1st IEEE Signal Processing Workshop Wireless Communications,* Paris, France, Apr. 1997, pp. 121–124.
[31] H. Liu and G. B. Giannakis, "Deterministic approaches for blind equalization of time-varying channels with antenna arrays," *IEEE Trans. Signal Processing,* to appear.
[32] H. Liu, G. B. Giannakis, and M. K. Tsatsanis, "Time-varying system identification: A deterministic blind approach using antenna arrays," in *Proc. of 30th Conf. Information Science Systems,* Princeton Univ., Princeton, NJ, Mar. 20–22, 1996, pp. 880–884.
[33] 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.
[34] B. Porat and B. Friedlander, "Blind equalization of digital communication channels using higher-order statistics," *IEEE Trans. Signal Processing*, vol. 39, pp. 522–526, 1991.
[35] J. Proakis, *Digital Communications.* New York: McGraw Hill, 1989.
[36] ——, "Adaptive equalization techniques for acoustic telemetry channels," *IEEE J. Oceanic Eng.,* vol. 16, pp. 21–31, Jan. 1991.
[37] T. S. Rappaport, S. Y. Seidel, and R. Singh, "900-MHz multipath propagation measurements for U.S. digital cellular radiotelephone," *IEEE Trans. Veh. Technol.*, vol. 39, pp. 132–139, May 1990.
[38] A. W. Rihaczek, *Principles of High-Resolution Radar.* CA: Peninsula, 1985.
[39] A. M. Sayeed and B. Aazhang, "Exploiting Doppler diversity in mobile wireless communications," in *Proc. 31st Conf. Information Sciences Systems*, The Johns Hopkins Univ., Baltimore, MD, Mar. 19–21, 1997, vol. I, pp. 287–292.
[40] A. Scaglione, S. Babarossa, 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.
[41] E. Serpedin and G. B. Giannakis, "Blind channel identification and equalization with modulation induced cyclostationarity," in *Proc. 31st Conf. Info. Sciences Systems,* The Johns Hopkins Univ., Baltimore, MD, Mar. 19–21, 1997, vol. II, pp. 792–797.
[42] R. Sharma, W. D. Grover, and W. A. Krzymien, "Forward-error-control (FEC)-assisted adaptive equalization for digital cellular mobile radio," *IEEE Trans. Veh. Technol.,* vol. 42, pp. 94–102, Feb. 1993.
[43] D. T. M. Slock, "Blind fractionally-spaced equalization, perfect-reconstruction filter banks and multichannel linear prediction," in *Proc. Int. Conf. ASSP,* Adelaide, Australia, 1994, vol. IV, pp. 585–588.
[44] M. Stojanovic, J. Catipovic, and J. G. Proakis, "Adaptive multichannel combining and equalization for underwater acoustic communications," *J. Acoustical Soc. Amer.,* pp. 1621–1631, Sept. 1993.
[45] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inform. Theory,* pp. 340–349, Mar. 1994.
[46] M. K. Tsatsanis and G. B. Giannakis, "Modeling and equalization of rapidly fading channels," *Intl. J. Adaptive Contr. Signal Processing,* vol. 10, pp. 159–176, Mar. 1996.
[47] ——, "Equalization of rapidly fading channels: Self recovering methods," *IEEE Trans. Commun.,* vol. 44, pp. 619–630, May 1996.

[48] ——, "Transmitter induced cyclostationarity for blind channel equalization," *IEEE Trans. Signal Processing,* vol. 45, pp. 1785–1794, July 1997.

[49] ——, "Subspace methods for blind estimation of time-varying channels," *IEEE Trans. Signal Processing,* vol. 45, pp. 3084–3093, Dec. 1997.

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

[51] J. K. Tugnait, "Identification of linear stochastic systems via second- and fourth-order cumulant matching," *IEEE Trans. Inform. Theory,* vol. 33, pp. 393–407, May 1987.

[52] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Trans. Signal Processing*, vol. 43, pp. 2982–2993, Dec. 1995.

Georgios B. Giannakis (Fellow, IEEE) received the Diploma in electrical engineering from the National Technical University of Athens, Greece, in 1981. He received the M.Sc. degrees in electrical engineering and mathematics in 1983 and 1986, respectively, and the Ph.D. degree in electrical engineering in 1986 from the the University of Southern California (USC), Los Angeles, CA.

After lecturing for one year at USC, he joined the University of Virginia, Charlottesville, in September 1987, where he is now a Professor with the Department of Electrical Engineering. His general interests are in the areas of signal processing, communications, estimation and detection theory, and system identification.

Dr. Giannakis received the IEEE Signal Processing Society's 1992 Paper Award in the Statistical Signal and Array Processing (SSAP) area. 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 was a Guest Editor for two special issues on high-order statistics for the *International Journal of Adaptive Control and Signal Processing,* for the EURASIP journal *Signal Processing,* and for the IEEE TRANSACTIONS ON SIGNAL PROCESSING January 1997 Special Issue on Signal Processing for Advanced Communications. He has served as an Associate Editor for IEEE TRANSACTIONS ON SIGNAL PROCESSING and IEEE SIGNAL PROCESSING LETTERS, as well as Secretary of the Signal Processing Conference Board. He has been a member of the SP Publications board, the IMS, the European Association for Signal Processing, Vice-Chair the SSAP Technical Committee, and Chair of the SP Technical Committee for Communications.

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 1995 he has been working as a Teaching and Research Assistant at University of Virginia and is currently a Ph.D. student in the CCSP Labaratory. His research interests include statistical signal processing, system identification, time-varying systems, and equalization of fading channels in digital communications.