

Pilot Symbol Assisted Modulation in Frequency Selective Fading Wireless Channels

Michail K. Tsatsanis, *Member, IEEE*, and Zhengyuan (Daniel) Xu, *Member, IEEE*

Abstract—Frequency-selective fading channels are typically modeled either as a combination of Doppler components or as lowpass stochastic processes. In both cases, accurate parameter and/or Doppler frequency estimation is impeded by the fact that the Doppler frequencies are typically very low (compared with the data rate) and closely spaced. This problem is mitigated in pilot symbol assisted modulation (PSAM) systems that employ distributed training. Those systems can provide information about a time-undersampled version of the channel that may be easier to identify. In this paper, we address the problem of estimating the fading channel's correlation matrices from the received data by exploiting the distributed training symbols. Multichannel autoregressive (AR) models are estimated to fit the channel's variations, and the Doppler frequencies are identified through the peaks of the AR spectrum. The performance of the proposed methods is studied through analytical and experimental results. Finally, Kalman filtering ideas are employed to track the time-varying channel taps based on the estimated AR model.

Index Terms—Frequency-selective fading, Kalman filtering, PSAM.

I. INTRODUCTION

ESTIMATION of fading channels is a problem of considerable importance in wireless communication setups when strong Doppler shifts are present. Applications include communications with high-speed vehicles, microwave links, or underwater acoustic channels [3]. Apart from the clearly necessary task of tracking the channel variations, recent studies have shown that *long-term prediction* of the channel gains can be exploited in a number of beneficial ways when available. For example, if a feedback channel is available, channel predictions can be used for power control purposes [10], [9] and adaptive coding [13], [14].

In most wireless scenarios, a preamble of training symbols is used to aid the receiver in identifying and equalizing the channel. In the case of fast fading channels, however, we will find it useful to consider distributed training, i.e., the insertion of one training symbol every P transmitted symbols. This technique is also known as pilot symbol assisted modulation

(PSAM) (e.g., [25]). In this paper, we study channel estimation issues for PSAM systems.

Fading channels have been studied extensively, and several models have been developed to describe their time variations. Analysis of PSAM schemes in flat Rayleigh fading channels is provided by [4] and for frequency selective channels in [5]. For both, it is discussed in [21] that coherent demodulation can be achieved from pilot-symbol aided estimation of the fading channel characteristics. Other contributions on fading channel compensation under the PSAM scheme can be found in [1], [27], and [32]. In many cases, the channel taps are modeled as general lowpass stochastic processes (e.g., [17]), the statistics of which depend on mobility parameters. A different approach explicitly models the channel response by the amplitudes and delays of each multipath component (e.g., [18], [31]). Such models, however, have, until recently, not been incorporated in the development of equalization algorithms. Typically, standard adaptive algorithms are used in fading environments, and the time-varying (TV) channel models are only exploited to assess the performance of the algorithms.

The reason for this approach lies in the difficulty of estimating the parameters of the channel variation model since the channel taps are not observed directly. Recent progress, however, in this area has been reported. Suboptimal solutions to the problem of estimating the statistics of a random tap channel and fitting appropriate models were proposed in [28], whereas optimal maximum likelihood solutions were studied in [8]. In [7], a combination of two Kalman filters was utilized to track the input and channel, respectively. Estimation issues in the context of explicit parameterization by time-varying amplitudes and delays were studied in [29] and [30] for a single antenna and in [22] for multiple antennas. Optimal solutions were reported in [6]. The distributed training case was studied in [11] for flat fading channels.

An added difficulty to fitting a model to the channel's tap variations is the fact that those variations are relatively slow (compared with the bit rate). Therefore, they possess spectral components of very low frequencies that are difficult to discriminate or estimate with precision. For this reason, distributed training appears to be naturally suitable for the current setup.

In this paper, we model the TV channel response vector as a multichannel autoregressive (AR) process, and we develop an algorithm to estimate the AR parameters. The proposed AR modeling scheme is fairly flexible. It can be thought of as an approximation to the model of [17] (see also [11]) or as a model for channel tap trajectories consisting of distinct Doppler components [29]). The proposed method is free of convergence problems and does not suffer from the limitations of [8] and [28].

Manuscript received July 12, 1999; revised March 31, 2000. This work was supported by NSF/CAREER CCR-9733048, NSF/NCR-9706658, NSF/Wireless CCR-9979288, and NJCST 99-2042-007-17. The associate editor coordinating the review of this paper and approving it for publication was Prof. Colin F. N. Cowan.

M. K. Tsatsanis is with the Electrical and Computer Engineering Department, Stevens Institute of Technology, Hoboken, NJ 07030 USA (e-mail: mtatsan@stevens-tech.edu).

Z. Xu is with the Department of Electrical Engineering, University of California, Riverside, CA 92521 USA (e-mail: dxu@ee.ucr.edu).

Publisher Item Identifier S 1053-587X(00)05972-9.

Once the model parameters are obtained, Kalman filtering techniques may be applied to track the channel taps.

The rest of the paper is organized as follows: A discrete-time input/output model for a wireless communication system with multipath fading channels is presented in Section II. Section III associates modeling of the fading channel with a multirate AR process. The proposed algorithms to estimate the AR model parameters and correlations of the fading channel taps are developed in Section IV, whereas Doppler shift estimation is briefly discussed in Section V. The performance of the proposed method is studied in Section VI, and a channel tracking algorithm is given in Section VII. Finally, some simulation results are presented in Section VIII, and conclusions are made in Section IX.

II. PROBLEM STATEMENT

Let us consider a digital wireless transmission with and i.i.d. input $w(n)$ that takes values from a finite constellation. Suppose that $w(n)$ is linearly modulated using a pulse¹ $g(t)$ and is received through a time-varying multipath channel with L distinct paths with gains $\alpha_i(t)$, $i = 1, \dots, L$. If we sample the baseband received signal $y(t)$ at every symbol period T_s , then the received discrete time signal becomes

$$\begin{aligned} y(n) &\stackrel{\text{def}}{=} y(t)|_{t=nT_s} \\ &= \sum_{k=-\infty}^{\infty} w(k) \sum_{i=1}^L \alpha_i(nT_s) g((n-k)T_s - \tau_i) + v(n) \end{aligned} \quad (1)$$

or equivalently

$$y(n) = \sum_{k=-\infty}^{\infty} w(n-k) \sum_{i=1}^L \alpha_i(nT_s) g(kT_s - \tau_i) + v(n) \quad (2)$$

where τ_i is the path's delay, and $v(n)$ is AWGN with variance σ_v^2 .

When the transmitted signal propagates through an unknown channel, it is customary to utilize fractional sampling at the receiver to ensure that sufficient statistics are retained. In the current framework, the model of (2) can be easily extended to multiple antennas and/or fractional sampling by modifying $y(n)$ and $g(kT_s - \tau_i)$ (e.g., [22] and [23]). However, we have chosen to present the proposed methods using the model of (2) in the interest of clarity of presentation. By defining the TV channel $h(n-k; k)$ as

$$h(n-k; k) \stackrel{\text{def}}{=} \mathbf{g}^T(k) \boldsymbol{\alpha}(n) \quad (3)$$

with

$$\begin{aligned} \mathbf{g}(k) &= [g(kT_s - \tau_1) \cdots g(kT_s - \tau_L)]^T \\ \boldsymbol{\alpha}(n) &= [\alpha_1(nT_s) \cdots \alpha_L(nT_s)]^T \end{aligned} \quad (4)$$

¹ $g(t)$ includes all time-invariant filtering occurring either at the transmitter or receiver.

and assuming it has order q , the input/output (I/O) relationship (2) can be expressed by

$$y(n) = \sum_{k=0}^q h(n-k; k) w(n-k) + v(n). \quad (5)$$

In order to track the fading channel, a pilot symbol is sent as training every P bits. If we collect $M+1$ successive received data points in a vector $\mathbf{y}(n) = [y(n) \cdots y(n-M)]^T$, then the vector form of (5) is

$$\mathbf{y}(n) = \mathbf{H}(n) \mathbf{w}(n) + \mathbf{v}(n) \quad (6)$$

where

$$\begin{aligned} \mathbf{H}(n) &= \begin{bmatrix} h(n; 0) & \cdots & h(n-q; q) & & 0 \\ & \ddots & & \ddots & \\ 0 & & h(n-M; 0) & \cdots & h(n-M-q; q) \end{bmatrix} \\ \mathbf{w}(n) &= [w(n) \quad w(n-1) \quad \cdots \quad w(n-M-q)]^T \\ \mathbf{v}(n) &= [v(n) \quad v(n-1) \quad \cdots \quad v(n-M)]^T. \end{aligned} \quad (7)$$

The input/output (I/O) relationship of (6) has too many degrees of freedom and does not provide a useful model, unless some assumptions are made on the variation of $h(n-k; k)$ as a function of n . Otherwise, if $h(n-k; k)$ is allowed to vary arbitrarily, (6) involves $q+1$ different channel parameters for every received data point $y(n)$.

Since the channel variations are due to Doppler shifts, a simple model may be to consider $h(n; k)$ as a combination of Doppler components (e.g., [11] and [18])

$$h(n; k) = \sum_{l=1}^L \theta_{kl} e^{j\omega_l n} + \varepsilon(n; k) \quad (8)$$

where L is the total number of paths, and ω_l is the Doppler shift at each path; the noise $\varepsilon(n; k)$ accounts for modeling errors. The model of (8) is accurate when there is a small number of well defined, distinct paths but may not be useful for the case of transmissions through a random dispersive medium as, for example, in underwater acoustic channels [3]. In this case, the spectral components of the channel's variations are better captured by modeling $h(n-k; k)$ as a random process [12], [16], [17]. AR approximations have been proposed [11], [28] in which case, the random channel framework can also approximate the harmonics-in-noise model of (8) (see also [18]).

Unfortunately, the accurate estimation of the AR parameters (or the frequencies ω_l) is impeded by the fact that the Doppler frequencies are relatively low (compared with the data rate) and closely spaced, and hence, $h(n-k; k)$ is a highly lowpass process. A long training record would therefore be required to identify the model. This problem is avoided in PSAM systems since the training symbols are dispersed, and their response corresponds to an undersampled version of $h(n-k; k)$, as explained next.

Considering (6), and in order that the signature of the pilot symbol $w(nP)$ be included in the output vector for our estimation purposes, we focus on a data window in the time interval $(nP+d-M, nP+d)$ for some offset d , i.e., we consider $\mathbf{y}(nP+d)$ for $n = 1, \dots, N$.

Notice that $w(nP)$ is the $(d+1)$ th element of $\mathbf{w}(nP+d)$ [cf. (7)]. If we write $\mathbf{H}(nP+d)$ in (7) as

$$\mathbf{H}(nP+d) = [h_0(nP+d), \dots, \mathbf{h}_{M+q}(nP+d-M-q)] \quad (9)$$

with the $(i+1)$ th column denoted as $h_i(nP+d-i)$ corresponding to the signature of the bit $w(nP+d-i)$, then the signature of the training bit $w(nP)$ is the $(d+1)$ th column of $\mathbf{H}(nP+d)$, i.e., $h_d(nP)$. If $d \geq q+1$ and $M+1 \geq d$, then the pilot signature $h_d(nP)$ has the following particular form [see (6) and (7)]:

$$h_d(nP) = [0 \cdots 0 \ \mathbf{h}^T(nP) \ 0 \cdots 0]^T \quad (10)$$

which contains all the channel taps at time instant nP in a smaller size channel vector $\mathbf{h}(nP)$

$$\mathbf{h}(nP) = [h(nP; q) \cdots h(nP; 0)]^T \quad (11)$$

padding by leading and trailing zeros. Equation (6) now becomes

$$\begin{aligned} \mathbf{y}(nP+d) &= h_d(nP)w(nP) + \sum_{\substack{m=0 \\ m \neq d}}^{M+q} w(nP+d-m) \\ &\quad \cdot h_m(nP+d-m) + \mathbf{v}(nP+d). \end{aligned} \quad (12)$$

The signature $\mathbf{h}_d(nP)$ contains spectral components at frequencies $\omega_l P$ and is easier to identify. According to (3) and (11), we have

$$\mathbf{h}(nP) = \begin{bmatrix} \mathbf{g}^T(q)\boldsymbol{\alpha}(nP+q) \\ \vdots \\ \mathbf{g}^T(0)\boldsymbol{\alpha}(nP) \end{bmatrix}. \quad (13)$$

Since our method is based on the second-order statistics of the channel, we will explore the relationship between the correlation matrix of the undersampled channel

$$\mathbf{R}_h(kP) = E\{\mathbf{h}(nP+kP)\mathbf{h}^H(nP)\} \quad (14)$$

and that of the path gain

$$\mathbf{R}_\alpha(kP) = E\{\boldsymbol{\alpha}(nP+kP)\boldsymbol{\alpha}^H(nP)\}.$$

We have shown that the fading process and the TV channel are related to each other by (13). For a moderate channel order q and slow fading, $\boldsymbol{\alpha}(n)$ is approximately constant for $q+1$ symbol periods. Therefore, (13) can be approximated by

$$\mathbf{h}(nP) \approx \mathbf{G}^T \boldsymbol{\alpha}(nP) \quad (15)$$

where

$$\mathbf{G} = [\mathbf{g}(q) \cdots \mathbf{g}(0)]. \quad (16)$$

Hence, from (14) and (15), we have

$$\mathbf{R}_h(kP) \approx \mathbf{G}^T \mathbf{R}_\alpha(kP) \mathbf{G}. \quad (17)$$

Following common practice, it is reasonable to consider Rayleigh fading (in the absence of line-of-sight) [25], [26]. The statistics of the channel's time variation depends on several mobility parameters, and various models may be applicable. For example, in the Jakes' model, the correlation with lag k

of the random gain for the l th path $\alpha_l(nT_s)$ is a zeroth-order Bessel function of the first kind (see [17] for details).

In this paper, we will present a method to estimate $\mathbf{R}_h(kP)$ directly from the received signal without stipulating a particular mobility model. According to (17), $\mathbf{R}_\alpha(kP)$ can be easily obtained from $\mathbf{R}_h(kP)$ as long as \mathbf{G} has full row rank, and the timing is available. It requires that the channel order and the total number of paths satisfy $q+1 \geq L$.

III. AR MODEL CONSIDERATIONS

Our method will be applied in an AR framework. We will thus approximate the fading model (e.g., Jakes' fading model) with a general AR(p) model of order p for the time under-sampled gains (see also [11])

$$\begin{aligned} \alpha_l(nP) + \sum_{m=1}^p \beta_{lm} \alpha_l(nP-mP) &= e_l(nP) \\ l &= 1, \dots, L \end{aligned} \quad (18)$$

where $e_l(nP)$ is a white Gaussian random process with variance σ_l^2 . The undersampling factor P should be chosen based on the maximum Doppler shifts for all L paths such that the under-sampled signal $\alpha_l(nP)$ for each path still satisfies the Nyquist sampling theorem.

Equation (18) can be written in a compact matrix form by collecting $\alpha_l(nP)$ for different paths in a vector $\boldsymbol{\alpha}(nP)$ and similarly putting all noise terms $e_l(nP)$ in $\mathbf{e}(nP)$

$$\boldsymbol{\alpha}(nP) + \sum_{m=1}^p \mathbf{B}_m \boldsymbol{\alpha}(nP-mP) = \mathbf{e}(nP) \quad (19)$$

with

$$\mathbf{B}_m = \text{diag}(\beta_{1m}, \dots, \beta_{Lm}).$$

According to (15) and (19), the undersampled overall channel now also satisfies a general AR(p) model

$$\mathbf{h}(nP) + \sum_{m=1}^p \mathbf{A}_m \mathbf{h}(nP-mP) = \mathbf{u}(nP) \quad (20)$$

where the new AR parameters \mathbf{A}_m are related to \mathbf{B}_m by the linear transformation

$$\mathbf{A}_m = \mathbf{G}^T \mathbf{B}_m \quad (21)$$

and $\mathbf{u}(nP)$ is the excitation noise. Although both (19) and (21) can be used for channel estimation, it is more convenient to estimate \mathbf{A}_m based on (20) if \mathbf{G} is unknown. However, the estimation of \mathbf{A}_m is still not trivial since the channel $\mathbf{h}(nP)$ is not observed directly. One needs to somehow acquire the channel correlations $\mathbf{R}_h(kP)$ in order to solve the Yule-Walker equations

$$\mathbf{R}_h(kP) = - \sum_{m=1}^p \mathbf{A}_m \mathbf{R}_h(kP-mP) + \delta(k) \boldsymbol{\Sigma}_u \quad (22)$$

and obtain \mathbf{A}_m , where $\boldsymbol{\Sigma}_u$ is the covariance matrix of $\mathbf{u}(nP)$.

Similarly, if the exponentials-in-noise model of (8) is adopted, the channel correlations are still needed to estimate

the frequencies ω_l . Indeed, several spectral analysis techniques may be employed to obtain ω_l once $\mathbf{R}_h(kP)$ is provided. We therefore focus next on the estimation of $\mathbf{R}_h(kP)$ in the current PSAM framework. Our approach will aim for simple suboptimal methods. For optimal maximum likelihood approaches, see [8].

IV. PROPOSED ALGORITHM

Our task is to estimate the correlation $\mathbf{R}_h(kP)$ in (14). Hence, further estimation of either AR parameters or Doppler shifts can be possibly pursued. From (10), it suffices to estimate the correlations of $h_d(nP)$ defined by

$$\mathcal{R}_{h_d}(kP) \stackrel{\text{def}}{=} E\{h_d(nP+kP)h_d^H(nP)\} \quad (23)$$

in this way being insensitive to possible channel order overestimation. The following explicit expression for $h_d(nP)$ holds according to the I/O relationship (12)

$$h_d(nP) = \frac{1}{w(nP)} \left[\mathbf{y}(nP+d) - \sum_{\substack{m=0 \\ m \neq d}}^{M+q} w(nP+d-m) \cdot h_m(nP+d-m) - \mathbf{v}(nP+d) \right]. \quad (24)$$

The right-hand side of (24) consists of three terms. If we substitute (24) into (23) (and under certain length conditions), most cross-terms are zero for $k \neq 0$ due to the i.i.d. nature of $w(n)$. Then, (23) reduces to a data-correlation

$$\mathcal{R}_{h_d}(kP) = E \left\{ \frac{\mathbf{y}(nP+kP+d)\mathbf{y}^H(nP+d)}{w(nP+kP)w^*(nP)} \right\}. \quad (25)$$

Equation (25) is valid if the signature of the nP th pilot symbol does not overlap with the signature of the $(nP+kP)$ th pilot symbol, i.e., when $P > M+q$. Furthermore, in the case $k=0$, signatures within the same block are correlated with themselves, and thus, the channel correlation $\mathcal{R}_{h_d}(0)$ requires separate treatment, as discussed later in this section. Then, from (25), for given $n=1, \dots, N$ observation windows of received data, we can obtain an estimator $\hat{\mathcal{R}}_{h_d}(kP)$ for $k > 0$

$$\hat{\mathcal{R}}_{h_d}(kP) = \frac{1}{N-k} \sum_{n=1}^{N-k} \frac{\mathbf{y}(nP+kP+d)\mathbf{y}^H(nP+d)}{w(nP+kP)w^*(nP)} \quad (26)$$

provided that $w(nP) \neq 0$. From (10) and (23), it is easy to show that the correlation function for the fading channel $\mathbf{R}_h(kP)$ is related to $\mathcal{R}_{h_d}(kP)$ as follows:

$$\mathbf{R}_h(kP) = \mathbf{Q}\mathcal{R}_{h_d}(kP)\mathbf{Q}^T \quad (27)$$

where $\mathbf{Q} = [\mathbf{0} \ \mathbf{I}_{q+1} \ \mathbf{0}]$ is a selection matrix, and \mathbf{I}_{q+1} is the identity matrix. Equation (27) shows that the proposed channel

correlation estimation is not sensitive to overestimation of the channel order q .

Equations (26) and (27) provide a method to estimate the correlation function with any nonzero lag k for the fading channel directly from the received data and are the fundamental relationships we will exploit next in the development of our algorithm.

It appears that the correlation estimates from (26) can be used directly in (22) to solve for \mathbf{A}_m . However, (26) is only correct for $k \neq 0$; hence, we need to avoid $\mathbf{R}_h(0)$ in (22). For simplicity, we denote $\mathbf{R}_h(kP)$ as $\mathbf{R}(k)$. If we first let $k = p+1, \dots, 2p$ in (22), we arrive at p equations to solve for $\mathbf{A}_1, \dots, \mathbf{A}_p$

$$[\mathbf{A}_1 \ \dots \ \mathbf{A}_p] \begin{bmatrix} \mathbf{R}(p) & \dots & \mathbf{R}(2p-1) \\ \mathbf{R}(p-1) & \ddots & \mathbf{R}(2p-2) \\ \vdots & & \vdots \\ \mathbf{R}(1) & \dots & \mathbf{R}(p) \end{bmatrix} = -[\mathbf{R}(p+1) \ \dots \ \mathbf{R}(2p)]. \quad (28)$$

Once these AR parameters are obtained, $\mathbf{R}(0)$ can also be estimated. If we set $k = 1, \dots, p$ in (22), another set of equations can be obtained to solve for $\mathbf{R}(0)$ by utilizing the estimated $\mathbf{A}_1, \dots, \mathbf{A}_p$ and $\mathbf{R}(1), \dots, \mathbf{R}(2p)$ up to this point

$$[\mathbf{A}_1^T \ \dots \ \mathbf{A}_p^T]^T \mathbf{R}(0) = - \begin{bmatrix} \mathbf{R}(1) + \mathbf{A}_2 \mathbf{R}^H(1) + \dots + \mathbf{A}_p \mathbf{R}^H(p-1) \\ \vdots \\ \mathbf{R}(p) + \mathbf{A}_1 \mathbf{R}(p-1) + \dots + \mathbf{A}_{p-1} \mathbf{R}(1) \end{bmatrix} \quad (29)$$

where the property of $\mathbf{R}(-k) = \mathbf{R}^H(k)$ has been used according to its definition.

Furthermore, the estimation for $\mathbf{A}_1, \dots, \mathbf{A}_p$ could be improved if we employ all the results we have obtained. Letting $k = 1, \dots, 2p$ in (22), we arrive at a new set of overdetermined equations that have p unknown matrices $\mathbf{A}_1, \dots, \mathbf{A}_p$ but $2p$ equations. By solving these equations, $\mathbf{A}_1, \dots, \mathbf{A}_p$ can be estimated. The proposed AR modeling procedure is therefore a two-step one and is summarized in Table I. This procedure may be further repeated to improve estimation accuracy.

A. Estimation of $\mathbf{R}(0)$

In the previous section, an AR modeling and spectral estimation method was presented to describe the channel's variations. It is clear that several other spectral estimation techniques may also be applicable once the correlations $\mathbf{R}(k)$ are available. For example, subspace methods of the MUSIC or ESPRIT type (e.g. [19]) could be used based on the model of (8). However, the estimator of (26) is not valid for $k=0$, and we have not yet provided a method of estimating $\mathbf{R}(0)$ from the data. This is a crucial detail on which the application of general spectral analysis techniques depends.

In order to derive an estimation method for $\mathbf{R}(0)$, let us consider the output correlation given pilot symbols $w(nP)$

$$\mathcal{R}_{y,0}(nP+d) = E\{\mathbf{y}(nP+d)\mathbf{y}^H(nP+d)\}. \quad (30)$$

TABLE I
 PROPOSED ALGORITHM

-
1. Estimate $\mathcal{R}_{h_d}(kP)$ for $k = 1, \dots, 2p$ based on (26)
 2. Operate on $\mathcal{R}_{h_d}(kP)$ as in (27) to obtain $\mathbf{R}_h(kP)$
 (if the channel order has been overestimated)
 3. Solve (28) for initial estimation of AR parameters
 4. Use the results in step 2 and 3 to solve (29) for $\mathbf{R}(0)$
 5. Use $\mathbf{R}_h(kP)$ for $k = 0, \dots, 2p$ to solve for
 improved estimation of AR parameters.
-

Let us denote by σ_{tr}^2 and σ_w^2 the power for the pilot symbols and the unknown transmitted symbols, respectively. Then, based on (12) and (30), $\mathcal{R}_{y,0}(nP+d)$ yields²

$$\mathcal{R}_{y,0} = \sigma_{\text{tr}}^2 E\{h_d(nP)h_d^H(nP)\} + \mathcal{R}_{\text{int}} \quad (31)$$

where \mathcal{R}_{int} accounts for the contribution of interfering symbols and noise

$$\begin{aligned} \mathcal{R}_{\text{int}} = & \sigma_w^2 \sum_{\substack{m=0 \\ m \neq d}}^{M+q} E\{h_m(nP+d-m)h_m^H(nP+d-m)\} \\ & + \sigma_v^2 \mathbf{I}_{M+1}. \end{aligned} \quad (32)$$

According to the structure of $\mathbf{H}(nP+d)$ [c.f. (7), (9), and (10)], it is easily seen that its $(m+1)$ th column $h_m(nP+d-m)$ is related to the channel vector $\mathbf{h}(nP+d-m)$ by the following transformation:

$$h_m(nP+d-m) = \mathbf{J}_m \mathbf{h}(nP+d-m) \quad (33)$$

where $\mathbf{h}(\cdot)$ is defined as in (11), and \mathbf{J}_m is a shifting matrix of dimension $(M+1) \times (q+1)$ with ones in the $(q-m)$ th diagonal.³ Therefore, (32) can be expressed by $\mathbf{h}(nP+d-m)$ as

$$\begin{aligned} \mathcal{R}_{\text{int}} = & \sigma_w^2 \sum_{\substack{m=0 \\ m \neq d}}^{M+q} \mathbf{J}_m E\{\mathbf{h}(nP+d-m)\mathbf{h}^H(nP+d-m)\} \\ & \cdot \mathbf{J}_m^T + \sigma_v^2 \mathbf{I}_{M+1}. \end{aligned} \quad (34)$$

If we assume that $h(n-k; k)$ is a stationary random process, then we have

$$\begin{aligned} E\{\mathbf{h}(nP+d-m)\mathbf{h}^H(nP+d-m)\} = & \mathbf{R}(0) \\ \forall m \in [0, M+q]. \end{aligned} \quad (35)$$

Substituting (35) in (31) and (34), respectively, we obtain

$$\mathcal{R}_{y,0} = \sigma_{\text{tr}}^2 \mathbf{J}_d \mathbf{R}(0) \mathbf{J}_d^T + \mathcal{R}_{\text{int}} \quad (36)$$

²We will drop the index “ $nP+d$ ” later because of the equal power for all pilot symbols. However, σ_{tr}^2 need not be equal to σ_w^2 .

³We can see that \mathbf{Q} in (27) equals \mathbf{J}_d .

$$\mathcal{R}_{\text{int}} = \sigma_w^2 \sum_{\substack{m=0 \\ m \neq d}}^{M+q} \mathbf{J}_m \mathbf{R}(0) \mathbf{J}_m^T + \sigma_v^2 \mathbf{I}_{M+1}. \quad (37)$$

Combining (36) and (37) together, we can express $\mathcal{R}_{y,0}$ by

$$\begin{aligned} \mathcal{R}_{y,0} = & (\sigma_{\text{tr}}^2 - \sigma_w^2) \mathbf{J}_d \mathbf{R}(0) \mathbf{J}_d^T + \sigma_w^2 \\ & \cdot \sum_{m=0}^{M+q} \mathbf{J}_m \mathbf{R}(0) \mathbf{J}_m^T + \sigma_v^2 \mathbf{I}_{M+1}. \end{aligned} \quad (38)$$

In (38), $\mathcal{R}_{y,0}$ can be estimated from the received data by the sample average of N available data vectors

$$\hat{\mathcal{R}}_{y,0} = \frac{1}{N} \sum_{n=1}^N \mathbf{y}(nP+d) \mathbf{y}^H(nP+d). \quad (39)$$

In our context, σ_{tr}^2 and σ_w^2 are assumed known. The only unknowns in (38) are $\mathbf{R}(0)$ and σ_v^2 . To solve for them, we introduce the “vec” operation (see [20, ch. 12]) and arrange them in a new vector

$$\mathbf{x} = \begin{bmatrix} \text{vec}(\mathbf{R}(0)) \\ \sigma_v^2 \end{bmatrix}. \quad (40)$$

By taking “vec” operation on both sides of (38), we obtain

$$\mathbf{r}_y = \mathbf{C} \mathbf{x} \quad (41)$$

where

$$\mathbf{r}_y = \text{vec}(\mathcal{R}_{y,0}) \quad (42)$$

$$\mathbf{C} = \begin{bmatrix} (\sigma_{\text{tr}}^2 - \sigma_w^2) \mathbf{J}_d \otimes \mathbf{J}_d + \sigma_w^2 \sum_{m=0}^{M+q} \mathbf{J}_m \otimes \mathbf{J}_m \text{vec}(\mathbf{I}_{M+1}) \\ \sigma_v^2 \end{bmatrix} \quad (43)$$

and \otimes represents the Kronecker product (see also [20, ch. 12]). Matrix \mathbf{C} is determined by the relative power of the pilot and the transmitted symbols. For a variety of choices for M, q , it is observed that this matrix is left invertible in most cases, except when $\sigma_{\text{tr}}^2 = \sigma_w^2$. Notice that \mathbf{C} does not depend on the data and can be studied *a priori*. Under the full rank condition of \mathbf{C} , \mathbf{x} can be estimated uniquely from (41) in the least square sense

$$\hat{\mathbf{x}} = \mathbf{C}^\dagger \hat{\mathbf{r}}_y, \quad \mathbf{C}^\dagger = (\mathbf{C}^H \mathbf{C})^{-1} \mathbf{C}^H, \quad \hat{\mathbf{r}}_y = \text{vec}(\hat{\mathcal{R}}_{y,0}). \quad (44)$$

Once \mathbf{x} is estimated, $\mathbf{R}(0)$ can also be obtained by the reverse operation of “vec” on the appropriate part of \mathbf{x} according to (40).

This method provides us with a direct estimate of $\mathbf{R}(0)$. Together with the estimates of $\mathbf{R}(k)$ for $k \neq 0$ discussed in the earlier part of the section, AR parameters can be directly obtained by solving the normal equations, thus simplifying our algorithm. Due to the additional information of the training symbols, the accuracy for estimates is expected to improve compared with the previous two-step procedure. This point will be verified by our simulation results.

V. DOPPLER SHIFT ESTIMATION

One of the applications of the proposed method is to estimate the Doppler shifts. This time, we assume the model (8) is locally valid for $k = 0, \dots, q$. We have presented an algorithm to estimate the correlations $\mathbf{R}(k)$ of the TV channel $h(nP; k)$. Then,

the associated AR parameters can be estimated from $\mathbf{R}(k)$ based on the proposed method. Following standard autoregressive spectral analysis techniques (e.g., [19]), one can plot the spectrum of this AR process, search the peaks corresponding to the multiple Doppler shifts $\omega_l P$, and find ω_l . A general discussion of multivariate AR spectrum can also be found in [24, ch. 9].

In order that the Nyquist sampling criterion be satisfied, the undersampling factor P should be chosen according to $(1/(PT_s)) \geq 2f_l$, where f_l is the Doppler frequency in hertz. In addition, considering the conditions discussed earlier [see discussions after (9) and (25)], we arrive at

$$\begin{aligned} M + q < P \leq \min \left\{ \frac{1}{2f_l T_s} \right\} \quad (l = 1, \dots, L) \\ q + 1 \leq d \leq M + 1 \end{aligned} \quad (45)$$

in order for the proposed method to be applicable.

VI. PERFORMANCE ANALYSIS

We have shown that the correlations of the undersampled TV channel coefficients with different lags can be estimated from received data by employing distributed pilot symbols at the input. In this section, we will investigate finite sample effects on the performance of the algorithm in terms of the mean and the variance of both the channel coefficients and their correlations.

Assume that we collect N observation vectors $\mathbf{y}(nP + d)$, $n = 1, \dots, N$. Recall that this output is explicitly expressed in (12) as a combination of three parts:

- 1) training;
- 2) interfering;
- 3) noise parts.

Under the assumption of zero-mean inputs as well as noise and given training bits $w(nP)$, the mean h_μ of $h_d(nP)$ can be computed from (24) as

$$h_\mu = E \left\{ \frac{1}{w(nP)} \mathbf{y}(nP + d) \right\}. \quad (46)$$

Therefore, it can be estimated by

$$\hat{h}_\mu = \frac{1}{N} \sum_{n=1}^N \frac{1}{w(nP)} \mathbf{y}(nP + d). \quad (47)$$

Substituting (12) in (47), we obtain

$$\hat{h}_\mu = \frac{1}{N} \sum_{n=1}^N (\mathbf{h}_d(nP) + \mathbf{a}(n)) \quad (48)$$

where

$$\mathbf{a}(n) = \frac{1}{w(nP)} \left[\begin{array}{c} \sum_{\substack{m=0 \\ m \neq d}}^{M+q} w(nP + d - m) \\ \cdot h_m(nP + d - m) + \mathbf{v}(nP + d) \end{array} \right]. \quad (49)$$

Equation (48) is our basis for the following analysis.

A. Estimation of the Mean h_μ

According to (48), if the inputs and noise are all zero-mean, we can immediately obtain the expected value of \hat{h}_μ , given pilot symbols

$$E \left\{ \hat{h}_\mu | w(nP) \right\} = h_\mu \quad (50)$$

which shows that \hat{h}_μ is an unbiased estimator of h_μ . For simplicity, we will assume that this mean is zero in the sequel. Then, using (48), we obtain

$$\begin{aligned} \text{cov}(\hat{h}_\mu) &\stackrel{\text{def}}{=} E \left\{ \hat{h}_\mu \hat{h}_\mu^H \right\} \\ &= \frac{1}{N^2} \sum_{n_1=1}^N \sum_{n_2=1}^N \mathcal{R}_{h_d}(n_1 P - n_2 P) \\ &\quad + \frac{1}{N^2} \sum_{n_1=1}^N \sum_{n_2=1}^N E \{ \mathbf{a}(n_1) \mathbf{a}^H(n_2) \} \end{aligned} \quad (51)$$

where

$$\mathcal{R}_{h_d}(n_1 P - n_2 P) = E \{ h_d(n_1 P) h_d^H(n_2 P) \}.$$

The first term in (51) can be transformed to a single summation similarly to [24, ch. 5]. In the second term, $E \{ \mathbf{a}(n_1) \mathbf{a}^H(n_2) \} = 0$ for $n_1 \neq n_2$, again due to the i.i.d. property of the inputs and noise. Therefore, (51) is reduced to

$$\begin{aligned} \text{cov}(\hat{h}_\mu) &= \frac{1}{N^2} \sum_{k=-N+1}^{N-1} (N - |k|) \mathcal{R}_{h_d}(kP) \\ &\quad + \frac{1}{N^2} \sum_{n=1}^N E \{ \mathbf{a}(n) \mathbf{a}^H(n) \}. \end{aligned} \quad (52)$$

Since $h_d(nP)$ is related to $\mathbf{h}(nP)$ by (33), their correlations $\mathcal{R}_{h_d}(kP)$, $\mathbf{R}(k)$ have the following correspondence:

$$\mathcal{R}_{h_d}(kP) = \mathbf{J}_d \mathbf{R}(k) \mathbf{J}_d^T. \quad (53)$$

From the definition of $\mathbf{a}(n)$ in (49) and \mathcal{R}_{int} in (32), it is evident for the second term that

$$E \{ \mathbf{a}(n) \mathbf{a}^H(n) \} = \frac{1}{\sigma_w^2} \mathcal{R}_{int}. \quad (54)$$

Hence, the covariance follows

$$\begin{aligned} \text{cov}(\hat{h}_\mu) &= \frac{1}{N} \left[\begin{array}{c} \sum_{k=-N+1}^{N-1} \left(1 - \frac{|k|}{N} \right) \mathbf{J}_d \mathbf{R}(k) \mathbf{J}_d^T + \frac{\sigma_w^2}{\sigma_{tr}^2} \\ \cdot \sum_{\substack{m=0 \\ m \neq d}}^{M+q} \mathbf{J}_m \mathbf{R}(0) \mathbf{J}_m^T + \frac{\sigma_v^2}{\sigma_{tr}^2} \mathbf{I}_{M+1} \end{array} \right]. \end{aligned} \quad (55)$$

If the AR process for the fading channel is stable, then the channel correlations $\|\mathbf{R}(k)\|$ are absolutely summable $\sum_{k=-\infty}^{\infty} \|\mathbf{R}(k)\| < \infty$ from which we can conclude that

$$\lim_{N \rightarrow \infty} \text{cov}(\hat{h}_\mu) = 0.$$

Therefore, \hat{h}_μ is a mean square consistent estimator of h_μ .

B. Estimation of the Correlations

Our estimation methods for $\mathbf{R}(k)$ depend on the lag k . Different methods are used for $k \neq 0$ and for $k = 0$. We will discuss them, respectively, next.

Case 1— $k \neq 0$: The estimate of $\mathcal{R}_{h_d}(kP)$ for $k > 0$ is given by (26). The case of $k < 0$ can be similarly dealt with. If we substitute (12) in it, we have

$$\begin{aligned} \hat{\mathcal{R}}_{h_d}(kP) &= \frac{1}{N-k} \sum_{n=1}^{N-k} [h_d(nP+kP) + \mathbf{a}(n+k)] \\ &\quad \cdot [h_d(nP) + \mathbf{a}(n)]^H. \end{aligned} \quad (56)$$

Taking expected value on both sides of (56), it can be easily shown that

$$E\{\hat{\mathcal{R}}_{h_d}(kP)\} = \mathcal{R}_{h_d}(kP)$$

because of properties of $w(n)$ and $v(n)$. Thus, $\hat{\mathcal{R}}_{h_d}(kP)$ is an unbiased estimator of $\mathcal{R}_{h_d}(kP)$. Next, we will evaluate the variance of this estimator. Let us introduce some notation at this point. We will use $\mathbf{x}^{(m)}(\cdot)$ or $\mathbf{x}^{(m)}$ to indicate the m th element of the vector $\mathbf{x}(\cdot)$ or \mathbf{x} , $\mathbf{X}^{(m,l)}(\cdot)$ to represent the (m, l) th element of the matrix $\mathbf{X}(\cdot)$. Let us further denote by $\hat{\gamma}_{i,j}(kP) = \hat{\mathcal{R}}_{h_d}^{(i,j)}(kP)$ and by

$$\begin{aligned} \chi(i, j) &= E\{[\hat{\gamma}_{i,j}(kP) - \gamma_{i,j}(kP)][\hat{\gamma}_{i,j}(kP) - \gamma_{i,j}(kP)]^*\} \\ &= E\{\hat{\gamma}_{i,j}(kP)\hat{\gamma}_{i,j}^*(kP) - \gamma_{i,j}(kP)\gamma_{i,j}^*(kP)\}. \end{aligned} \quad (57)$$

It is shown in Appendix A that $\chi(i, j)$ can be computed from the correlations of the channel with different lags as

$$\begin{aligned} \chi(i, j) &= \frac{1}{N-k} \sum_{t=-N+k+1}^{N-k-1} \left(1 - \frac{|t|}{N-k}\right) \\ &\quad \cdot \gamma_{i,i}(tP)\gamma_{j,j}^*(tP) + \frac{1}{(N-k)\sigma_{\text{tr}}^2} \\ &\quad \cdot \left[\gamma_{i,i}(0)\mathcal{R}_{\text{int}}(j, j) + \gamma_{j,j}^*(0)\mathcal{R}_{\text{int}}(i, i) \right. \\ &\quad \left. + \frac{1}{\sigma_{\text{tr}}^2}\mathcal{R}_{\text{int}}(i, i)\mathcal{R}_{\text{int}}(j, j) \right] \end{aligned} \quad (58)$$

where \mathcal{R}_{int} is given by (37). From (58), it can be concluded that the variance of the estimate of the (i, j) th element of $\mathcal{R}_{h_d}(kP)$ will converge to zero as $N \rightarrow \infty$ if $\sum_{t=-\infty}^{\infty} |\gamma_{i,i}(tP)\gamma_{j,j}^*(tP)| < \infty$, which establishes consistency.

Case 2— $k = 0$: When $k = 0$, $\mathbf{R}(0)$ is estimated based on (44). Then

$$E\{\hat{\mathbf{x}}\} = \mathbf{C}^\dagger E\{\hat{\mathbf{r}}_y\} = \mathbf{C}^\dagger \mathbf{r}_y.$$

Applying (41) and under the full-rank condition of \mathbf{C} , we have $E\{\hat{\mathbf{x}}\} = \mathbf{x}$. The estimate of $\mathbf{R}(0)$ is obtained by the reverse “vec” operation on $\hat{\mathbf{x}}$, and therefore, this also provides an unbiased estimator for $\mathbf{R}(0)$.

Next, we turn our attention to the covariance of our estimate $\hat{\mathbf{x}}$. From (44), this is given by the following matrix:

$$\text{COV}_{\hat{\mathbf{x}}} = \mathbf{C}^\dagger [E\{\hat{\mathbf{r}}_y \hat{\mathbf{r}}_y^H\} - \mathbf{r}_y \mathbf{r}_y^H] \mathbf{C}^{\dagger T}. \quad (59)$$

In (59), our estimate $\hat{\mathbf{r}}_y$ is obtained from sample averaging. For notational convenience, let us denote the received data vector $\mathbf{y}(nP+d)$ by \mathbf{y}_n . Then, according to (39) and (42), we arrive at

$$\hat{\mathbf{r}}_y = \frac{1}{N} \sum_{n=1}^N \text{vec}(\mathbf{y}_n \mathbf{y}_n^H).$$

Therefore, $E\{\hat{\mathbf{r}}_y \hat{\mathbf{r}}_y^H\}$ is readily computed by

$$E\{\hat{\mathbf{r}}_y \hat{\mathbf{r}}_y^H\} = \frac{1}{N^2} \sum_{n_1=1}^N \sum_{n_2=1}^N E\{\text{vec}(\mathbf{y}_{n_1} \mathbf{y}_{n_1}^H) \text{vec}^H(\mathbf{y}_{n_2} \mathbf{y}_{n_2}^H)\}. \quad (60)$$

Due to the stochastic nature of the channel parameters, no elegant relation can be derived between the fourth-order statistics of the output and the input. An alternative to this analytical intractability is to obtain estimates for (60) from the received data samples.

VII. CHANNEL TRACKING

Up to this point, we have discussed channel model identification issues, and we have presented techniques for estimating the channel correlations. Based on the correlations, the parameters of either the model of (20) or (8) can be estimated. Once the model parameters are identified, Kalman filtering ideas may be employed to track the time-varying channel coefficients.

An obstacle in developing this procedure is the fact that we have only modeled the time undersampled channel $\mathbf{h}(nP)$ in (20) and estimated only the undersampled correlation sequence $\mathbf{R}_h(kP)$. Since the Kalman filter would require a full-rate model, we should give some thought to the question of how the undersampled and the full-rate models are related.

We should clearly make the assumption here that the channel $\mathbf{h}(n)$ is a bandlimited process and that no aliasing is incurred by the undersampling by P . Therefore, the full correlation sequence $\mathbf{R}_h(k)$ may be recovered from the undersampled one $\mathbf{R}_h(kP)$ by interpolation.

Under this condition, the model of (8) is related to the undersampled one by simply changing the frequencies ω_l by a factor of P . The transformation of the random model of (20) to a full-rate one, however, is more involved. Notice that the

TABLE II
KALMAN-BASED, CHANNEL TRACKING ALGORITHM

Initialization:	$\hat{\mathbf{h}}(0 -1) = \mathbf{0}, \quad \mathbf{S}(0 -1) = \mathbf{0}$
Recursion:	$\hat{\mathbf{h}}(n+1 n) = [\mathbf{A} - \mathbf{K}(n)\mathbf{w}^T(n)]\hat{\mathbf{h}}(n n-1) + \mathbf{K}(n)y(n)$ $\mathbf{K}(n) = \mathcal{A}\mathbf{S}(n n-1)\mathbf{w}^T(n)[\mathbf{w}^T(n)\mathbf{S}(n n-1)\mathbf{w}^T(n) + \sigma_v^2]^{-1}$ $\mathbf{S}(n+1 n) = \mathcal{A}[\mathbf{S}(n n-1)$ $- \mathbf{S}(n n-1)\mathbf{w}^T(n)[\mathbf{w}^T(n)\mathbf{S}(n n-1)\mathbf{w}^T(n) + \sigma_v^2]^{-1}$ $\times \mathbf{w}^T(n)\mathbf{S}(n n-1)]\mathcal{A}^H + \sigma_u^2\mathbf{J}\mathbf{J}^H$

full-rate channel is bandlimited; hence, it cannot be exactly represented by an AR process (AR spectra are not bandlimited). An approximate AR spectral model can be obtained, however, by reducing the frequencies of the spectral peaks by a factor of $1/P$. A procedure to achieve that could be to find the poles of the undersampled model $Z_l = \rho_l e^{j\varphi_l}$, $l = 1, \dots, L$ and create the poles of the full-rate model by reducing the phase by $1/P$, i.e., $\tilde{Z}_l = \rho_l e^{j(\varphi_l/P)}$. Once the full-rate model parameters have been identified, Kalman filtering ideas may be employed to derive minimum variance estimators for the time-varying coefficients $h(n; k)$. The proposed method yields an adaptive full-rate channel tracking algorithm. However, it requires knowledge of $w(n)$, which is not available for all time instants. It can be obtained by a delayed or tentative decision, and therefore, the algorithm has to be implemented in a decision-feedback mode (see also [7] and [28]).

Let \mathbf{A}_m , $w(n)$, $y(n)$ be given, and let the determinant $|\mathbf{A}(z)|$ have zeros inside the unit circle, where $\mathbf{A}(z) = \mathbf{I} + \sum_{m=1}^p \mathbf{A}_m z^{-m}$. If we define $\underline{\mathbf{h}}(n) \stackrel{\text{def}}{=} [\mathbf{h}^T(n), \dots, \mathbf{h}^T(n-p+1)]^T$, then (20) can be written in state-space form as

$$\underline{\mathbf{h}}(n+1) = \begin{bmatrix} -\mathbf{A}_1 & -\mathbf{A}_2 & \cdots & -\mathbf{A}_p \\ \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ & \ddots & \ddots & \vdots \\ \mathbf{0} & & \mathbf{I} & \mathbf{0} \end{bmatrix} \underline{\mathbf{h}}(n) + \begin{bmatrix} \mathbf{u}(n) \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}. \quad (61)$$

Let \mathcal{A} be the block square matrix in the right-hand side of (61); then, the state equation becomes

$$\underline{\mathbf{h}}(n+1) = \mathcal{A}\underline{\mathbf{h}}(n) + \mathbf{J}\mathbf{u}(n) \quad (62)$$

where $\mathbf{J} \stackrel{\text{def}}{=} [\mathbf{I}, \mathbf{0}, \dots, \mathbf{0}]$. In order to obtain the measurement equation, define $\mathbf{w}(n) \stackrel{\text{def}}{=} [w(n), w(n-1), \dots, w(n-q)]^T$ and $\mathbf{w}(n) \stackrel{\text{def}}{=} [\mathbf{w}^T(n), \mathbf{0}^T, \dots, \mathbf{0}^T]^T$. Then, (5) can be written as

$$y(n) = \mathbf{w}^T(n)\underline{\mathbf{h}}(n) + v(n). \quad (63)$$

Equations (62) and (63) offer a state space representation of the fading channel model with transition matrix \mathcal{A} (which is assumed known in this section). Based on this representation, the minimum variance estimator for the state vector, i.e., the conditional expectation of $\underline{\mathbf{h}}(n)$, given $\{w(k), y(k)\}_{k=0}^{n-1}$, can

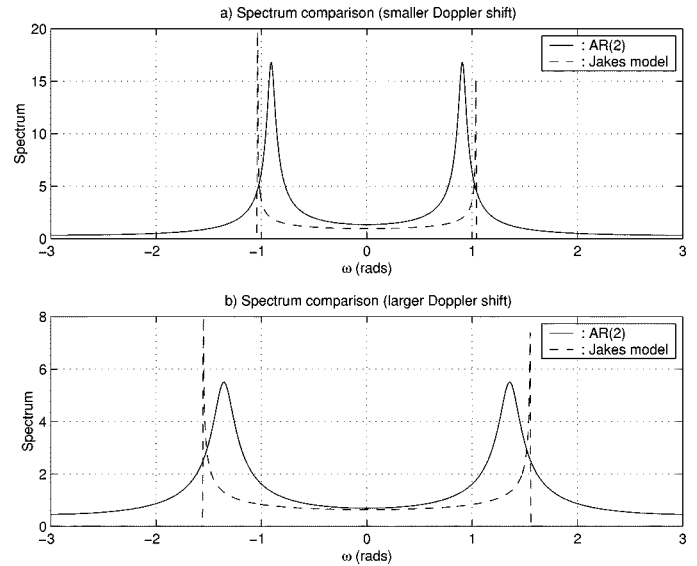


Fig. 1. Comparison of AR(2) spectrum with that of Jakes' model.

be computed using the Kalman filter [2]. The recursions are summarized in Table II (see also [2, p. 44]). Matrices $\mathbf{K}(n)$ and $\mathbf{S}(n|n)$ denote the Kalman gain and the covariance of the state vector $\underline{\mathbf{h}}(n)$, given data $\{w(i), y(i)\}_{i=0}^n$, respectively.

VIII. SIMULATIONS

We simulate a communication system with QPSK modulation. A raised-cosine pulse-shaping filter is chosen for $g(t)$ with roll-off factor $\alpha = 0.5$. The transmitted signal passes through two distinct paths. Each of them experiences independent Rayleigh fading with U-shape spectrum (see Jakes' model [17]) and normalized Doppler frequencies $\omega_1 = 0.0104$ rads and $\omega_2 = 0.0069$ rads. A training bit with double power is sent every $P = 150$ bit periods. An AR(2) model is used to approximate the Rayleigh fading process. These AR parameters are finally chosen for the two paths as follows: $\beta_{11} = -0.3809$, $\beta_{12} = 0.8139$ for the first path and $\beta_{21} = -1.1833$, $\beta_{22} = 0.9245$ for the second path. Our undersampled fading channel coefficients were then generated according to the AR(2) model in (19) with these parameters excited by complex AWGN with power 0.1. For the full-rate channel coefficients, we used linear interpolation based on those undersampled ones. We added 20 dB AWGN to the communication system. Other parameters are set to be $q = 1$, $d = 4$, and $M = 8$.

First, we compare the AR(2) model with Jakes' model in terms of their spectra to show the error introduced by our approximation. Fig. 1(a) shows the result for the smaller Doppler frequency ω_2 . The dashed line represents Jakes' model with $b_0 = 1$. The peaks of this spectrum occur at frequencies about $\pm\omega_2 P$. The solid line is for the AR(2) model. From this figure, we can observe the similar behavior of these two different channel models with some difference in the peak locations. The same conclusion can be made from Fig. 1(b) when Doppler frequency increases to ω_1 . However, this time, the frequency points corresponding to the peaks of the spectrum move to about $\omega = \pm\omega_2 P$ for both models.

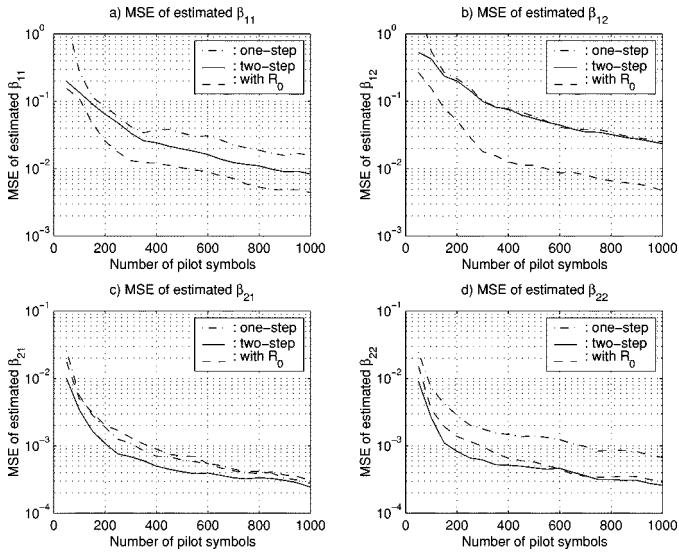
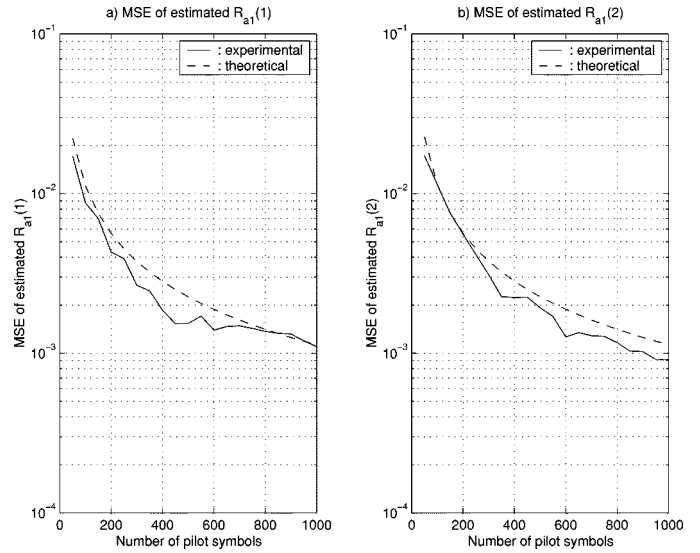
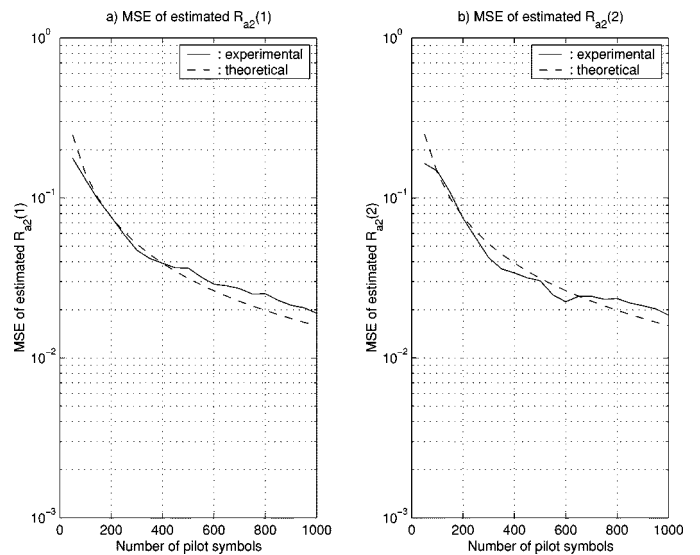


Fig. 2. MSE of estimated AR parameters.

Second, we estimate the AR model parameters β_{11} , β_{12} , β_{21} , and β_{22} by sending a total of 1000 pilot symbols. We ran 50 Monte Carlo iterations to obtain the average results. Fig. 2 shows the mean-square-error (MSE) of the estimates for β_{11} [in Fig. 2(a)], β_{12} [in Fig. 2(b)], β_{21} [in Fig. 2(c)], and β_{22} [in Fig. 2(d)] versus the number of available pilot symbols. The MSE's are measured in terms of the square of the error $E\{[\hat{\beta}_{i,j} - \beta_{i,j}]^2\}$ for $i, j = 1, 2$. Three different approaches to estimate the AR model parameters are studied. The first follows the one-step procedure without the knowledge of $\mathbf{R}(0)$, as described by (28). The second one is a two-step approach, as described in Table I. The third one employs the knowledge of estimated $\mathbf{R}(0)$ discussed in Section IV-A in estimating AR parameters. These results are plotted in the dashed-dotted line, the solid line, and the dashed line, respectively. It is clear that a small improvement on the accuracy of the estimates is achieved in the two-step approach. However, some improvement can also be observed by the direct estimation of $\mathbf{R}(0)$.

The direct application of our algorithm is to estimate the fading channel's statistics. Here, we also apply it to estimate $\mathbf{R}_\alpha(kP)$, which is the correlation matrix of the different path gains. In Fig. 3, the MSE of the correlation estimates for the gain $\alpha_1(nP)$ is shown for different number of pilot symbols [recall $\mathbf{R}_\alpha(kP)$ is diagonal]. These correlation estimates are obtained from the estimated $\mathbf{R}_h(kP)$ according to (17), assuming matrix \mathbf{G} and associated delays are known. The result for $k = 1$ is shown in Fig. 3(a) and for $k = 2$ in Fig. 3(b). For comparison, the theoretical MSE results according to (58) are also plotted in dashed lines, respectively. From this figure, the MSE's can be observed to decrease to 1×10^{-3} after 1000 pilot symbols are transmitted. In addition, good agreement is observed with the theoretically predicted performance. A similar conclusion can be drawn based on the results for the second path, as shown by Fig. 4(a) and (b). However, the error level is higher due to the larger value of true correlations.

Another application of the proposed method is to estimate the Doppler shifts based on the channel model (8) in our next experiment. We still use two Doppler frequencies: $\omega_1 = 0.0104$


 Fig. 3. MSE of estimated correlations for $\alpha_1(nP)$ with $k = 1$ and $k = 2$.

 Fig. 4. MSE of estimated correlations for $\alpha_2(nP)$ with $k = 1$ and $k = 2$.

rad/s and $\omega_2 = 0.0069$ rad/s. The coefficients θ_{11} , θ_{12} , θ_{21} , and θ_{22} are randomly generated in $[-1, 1]$. In total, 200 pilot symbols were available. Other simulation parameters were set as before. The channel correlations were first estimated based on our algorithm. An AR(2) model was fit to those correlations. Then, AR parameters can be estimated by three different approaches, as discussed before. The estimated Doppler frequencies were obtained using AR spectral analysis methods by peak-picking of the AR spectrum. In these three cases, their normalized MSE's are compared in Fig. 5(a) for the larger frequency and in Fig. 5(b) for the smaller one. Both of these figures show that the estimation error is larger if $\mathbf{R}(0)$ is not obtained in the first step, as indicated by dashed-dotted lines. An improvement can be achieved after $\mathbf{R}(0)$ is estimated in the second step, as presented by solid lines. Further improvement can be obtained by employing the estimated $\mathbf{R}(0)$ directly from the received signal, as shown by dashed lines.

In order to show the detailed performance for different iterations, we also plot the spectrum after 200 pilot symbols

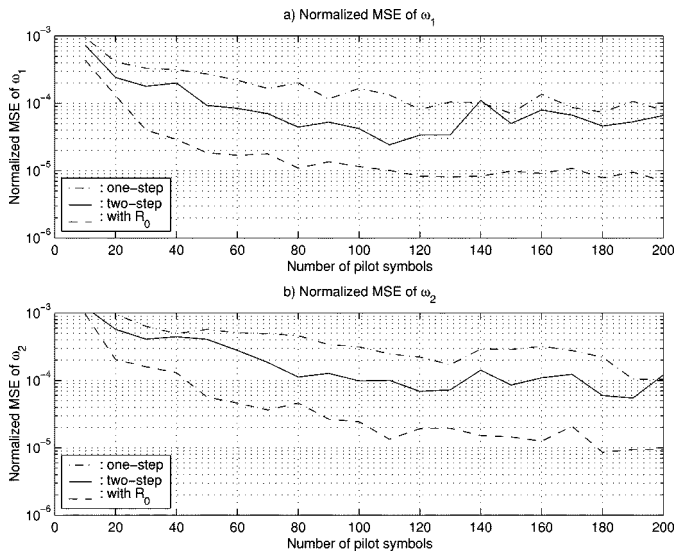


Fig. 5. Normalized MSE of estimated Doppler shifts.

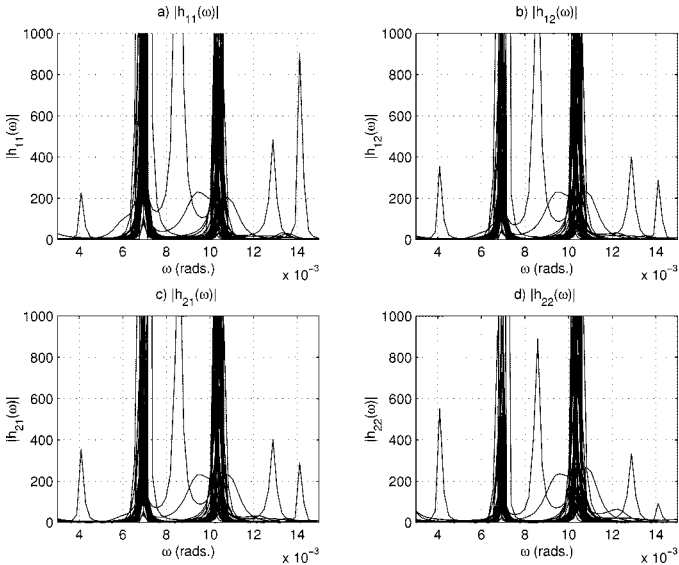


Fig. 6. AR spectrum from the one-step approach.

are transmitted. A 2×2 spectral matrix is obtained from estimated AR parameters according to [24, ch. 9] for 50 realizations. The four elements of the spectral matrix $\hat{\mathbf{H}}(\omega)$ ($|h_{11}(\omega)|$, $|h_{12}(\omega)|$, $|h_{21}(\omega)|$, $|h_{22}(\omega)|$) versus ω are shown in Fig. 6 for the first step estimation. For most of the realizations, peaks can be observed centering at the two true Doppler frequencies ω_1 and ω_2 , although some false peaks appear for a few iterations. This applies to all four elements. In the second step, even with estimated $\mathbf{R}(0)$, from those estimated AR parameters, the false peaks still exist in Fig. 7. However, after the more accurate estimate of $\mathbf{R}(0)$ is obtained, the spectrum estimate appears to be substantially more accurate, and false peaks disappear, as shown by Fig. 8. Therefore, the estimation error is expected to be reduced. This agrees with the simulation results at the point of 200 pilot symbols in the previous MSE figures, where the error level stays around 10^{-4} for the two-step procedure, while it reduces to 10^{-5} with the direct estimate of $\mathbf{R}(0)$.

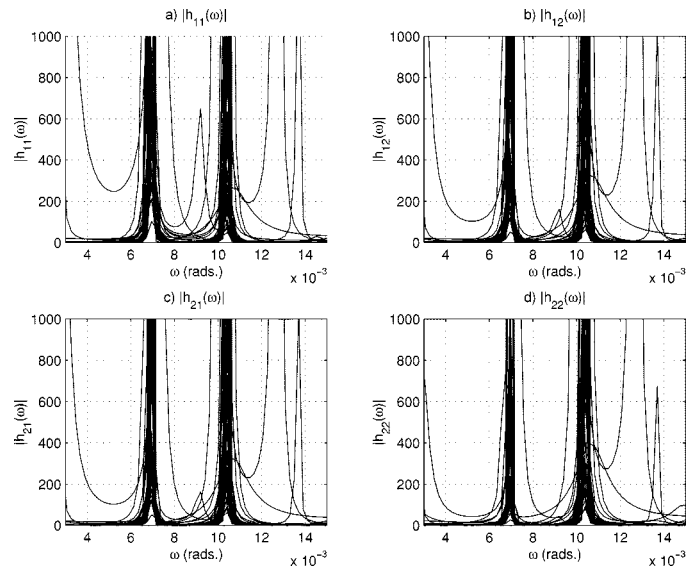
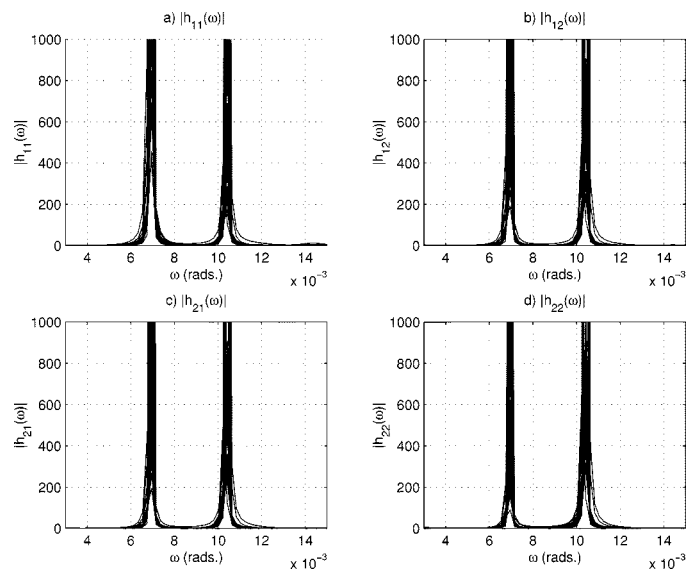


Fig. 7. AR spectrum from the two-step approach.

Fig. 8. AR spectrum with estimated $\mathbf{R}(0)$ from received data.

The proposed Kalman-based channel tracking algorithm based on pilot symbols only is compared with the channel tracking method with initial block training and decision feedback (DF) equalizer (e.g., [28]). The following AR(1) full-rate channel model is adopted:

$$\begin{bmatrix} h(n; 0) \\ h(n; 1) \end{bmatrix} = \begin{bmatrix} 0.98 & 0 \\ 0 & 0.95 \end{bmatrix} \begin{bmatrix} h(n-1; 0) \\ h(n-1; 1) \end{bmatrix} + \mathbf{u}(n)$$

where \mathbf{u} is circular complex AWGN with zero mean and variance 2×10^{-3} . The input 16-QAM signal is corrupted by 20 dB AWGN. In total, 2000 symbols are transmitted. In implementing [28], the initial 200 symbols are treated as training. After 200 symbol periods, the Kalman filter and DF(1,1) (both the feedforward and the feedback branch have order 1) are combined to track the fading channel and equalize the received symbols. To gain a reasonable comparison, we use a pilot symbol every tensymbols (10% of the transmission time) in our method.

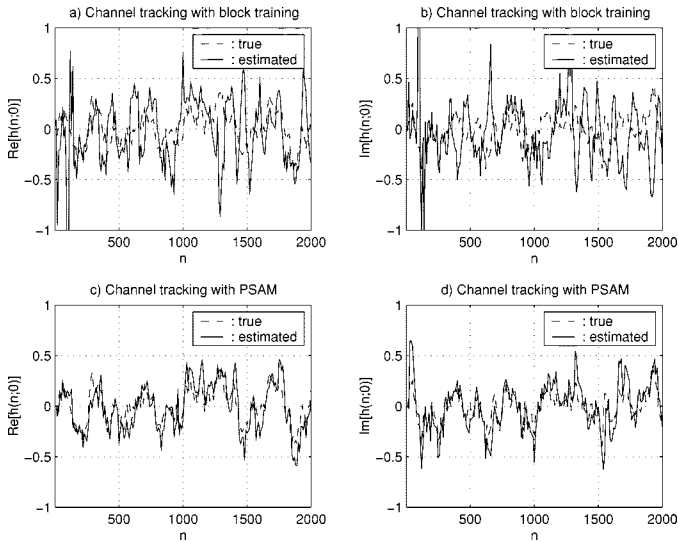


Fig. 9. Comparison of channel tracking performance between the proposed method (PSAM) and the method with block training and DFE

Since undersampled AR parameters can be estimated by the proposed method, the channel coefficients at those training time instants can be tracked based on pilot symbols. Then, the full-rate channel coefficients are obtained by linear interpolation from undersampled ones. From a typical realization, the tracking results for both real part (“Re”) and imaginary part (“Im”) of $h(n; 0)$ are obtained and shown by Fig. 9. All dashed lines in these figures represent the true channels. The channel tracking ability of our method (PSAM) can be observed from Fig. 9(c) and (d) in the current AR framework. With initial block training, [28] can track the fading channel in the training periods, as shown by Fig. 9(a) and (b). However, its tracking performance degrades after 200 symbols. Thus, a periodic retraining scheme is suggested in [28] to avoid the “run-away” effect caused by the DF equalizer.

In order to better evaluate the channel tracking performance of our method, we compare it with other previously developed PSAM-based methods in the presence of intersymbol interference (ISI). A typical method by Cavers [5] is adopted as a comparison. We assume two independent Rayleigh fading paths with equal power and a U-shape spectrum. The normalized cutoff frequency is set to be 0.005. The BPSK signal-to-noise ratio (SNR) is 20 dB. In every 10 bits, we assign the first bit as a training bit with 200 pilot bits available in total. The channel is assumed to have order 1. In implementing [5], we use $L = 5$, and thus, $K = 11$. The channel gain is estimated according to (13) in [5], where the optimal estimation coefficients are obtained by Wiener filter, as described in [4]. In our method, we use an AR(1) model to approximate the true undersampled channel variations. Our Kalman-based tracking procedure is same as in the previous experiment. The results for [5] are shown in Fig. 10(a) and (b), whereas the results are shown in Fig. 10(c) and (d) for our method. It can be seen that both methods can successfully track the channel variations and show equivalent performance. However, in implementing [5], we have assumed perfect knowledge of the channel correlations, whereas these correlations are completely unknown in our method and, thus, should be estimated for the estimation of

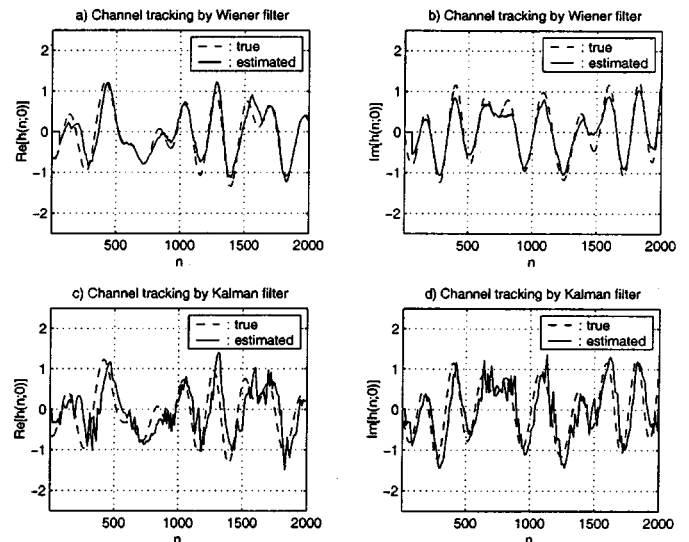


Fig. 10. Comparison of PSAM methods in channel tracking: the proposed Kalman-based method versus Wiener filter-based method.

AR parameters. Of course, more computational complexity is introduced along with this estimation process. It is expected that if the assumed channel correlations in [5] are appreciably different from the actual ones (model mismatch), the method will exhibit worse performance.

IX. CONCLUSIONS

The algorithms developed in this paper suggest that distributed training may be advantageous in identifying time-varying channels. We show that estimation of the channel correlations is possible in the PSAM framework with relatively simple algorithms. Furthermore, the estimators for the channel correlations are shown to be asymptotically mean square consistent. If the channel variations are approximated by an AR model, then AR parameters can also be estimated from channel correlations. Therefore, Kalman filtering ideas can be employed to track the TV channel. Moreover, based on the proposed algorithm, AR spectral analysis techniques are successfully demonstrated here for identifying the Doppler frequencies.

APPENDIX A DERIVATION OF (58)

To compute the first term in (57), we first write $\hat{\gamma}_{i,j}(kP)$ in an explicit form according to (56)

$$\hat{\gamma}_{i,j}(kP) = \frac{1}{N-k} \sum_{n=1}^{N-k} \left[h_d^{(i)}(nP + kP) + \mathbf{a}^{(i)}(n+k) \right] \cdot \left[h_d^{(j)}(nP) + \mathbf{a}^{(j)}(n) \right]^* \quad (64)$$

Substituting (64) in (57), we obtain

$$\chi(i,j) = \frac{1}{(N-k)^2} \sum_{n_1=1}^{N-k} \sum_{n_2=1}^{N-k} S(k, n_1, n_2) - \gamma_{i,j}(kP) \gamma_{i,j}^*(kP) \quad (65)$$

where

$$S(k, n_1, n_2) = E \left\{ \left[h_d^{(i)}(n_1P + kP) + \mathbf{a}^{(i)}(n_1 + k) \right] \cdot \left[h_d^{(j)}(n_1P) + \mathbf{a}^{(j)}(n_1) \right]^* \cdot \left[h_d^{(i)}(n_2P + kP) + \mathbf{a}^{(i)}(n_2 + k) \right]^* \cdot \left[h_d^{(j)}(n_2P) + \mathbf{a}^{(j)}(n_2) \right] \right\}. \quad (66)$$

If $h_d(nP)$ and $\mathbf{a}(n)$ were Gaussian, it can be shown that [24, ch. 5]

$$S(k, n_1, n_2) = S_1(k, n_1)S_2(k, n_2) + S_3(k, n_1, n_2)S_4(k, n_1, n_2)$$

where

$$S_1(k, n_1) = E \left\{ \left[h_d^{(i)}(n_1P + kP) + \mathbf{a}^{(i)}(n_1 + k) \right] \cdot \left[h_d^{(j)}(n_1P) + \mathbf{a}^{(j)}(n_1) \right]^* \right\}$$

$$S_2(k, n_2) = E \left\{ \left[h_d^{(i)}(n_2P + kP) + \mathbf{a}^{(i)}(n_2 + k) \right]^* \cdot \left[h_d^{(j)}(n_2P) + \mathbf{a}^{(j)}(n_2) \right] \right\}$$

$$S_3(k, n_1, n_2) = E \left\{ \left[h_d^{(i)}(n_1P + kP) + \mathbf{a}^{(i)}(n_1 + k) \right] \cdot \left[h_d^{(i)}(n_2P + kP) + \mathbf{a}^{(i)}(n_2 + k) \right]^* \right\}$$

$$S_4(k, n_1, n_2) = E \left\{ \left[h_d^{(j)}(n_1P) + \mathbf{a}^{(j)}(n_1) \right]^* \cdot \left[h_d^{(j)}(n_2P) + \mathbf{a}^{(j)}(n_2) \right] \right\}.$$

$h_d(nP)$ is indeed Gaussian due to our channel modeling assumptions. However, $\mathbf{a}(n)$ is given by (49) and can be shown to be Gaussian only if $w(n)$ is a constant modulus signal and the channel taps are circularly symmetric. Therefore, the rest of this analysis is restricted to a constant modulus $w(n)$. It is not hard to show that

$$S_1(k, n_1) = E \left\{ h_d^{(i)}(n_1P + kP) h_d^{*(j)}(n_1P) \right\} = \gamma_{i,j}(kP)$$

$$S_2(k, n_2) = E \left\{ \left[h_d^{(i)}(n_2P + kP) h_d^{*(j)}(n_2P) \right] \right\} = \gamma_{i,j}^*(kP)$$

and similarly

$$S_3(k, n_1, n_2) = \gamma_{i,i}(n_1P - n_2P) + E \left\{ \mathbf{a}^{(i)}(n_1 + k) \mathbf{a}^{*(i)}(n_2 + k) \right\}$$

$$S_4(k, n_1, n_2) = \gamma_{j,j}^*(n_1P - n_2P) + E \left\{ \mathbf{a}^{*(j)}(n_1) \mathbf{a}^{(j)}(n_2) \right\}.$$

Therefore, (65) is simplified to

$$\begin{aligned} \chi(i, j) &= \frac{1}{(N-k)^2} \sum_{n_1=1}^{N-k} \sum_{n_2=1}^{N-k} \\ &\cdot \left[\gamma_{i,i}(n_1P - n_2P) + E \left\{ \mathbf{a}^{(i)}(n_1 + k) \mathbf{a}^{*(i)}(n_2 + k) \right\} \right] \\ &\cdot \left[\gamma_{j,j}(n_1P - n_2P) + E \left\{ \mathbf{a}^{*(j)}(n_1) \mathbf{a}^{(j)}(n_2) \right\} \right]. \quad (67) \end{aligned}$$

In (67), $E\{\mathbf{a}^{(i)}(n_1 + k)\mathbf{a}^{*(i)}(n_2 + k)\}$, and $E\{\mathbf{a}^{*(j)}(n_1)\mathbf{a}^{(j)}(n_2)\}$ survive only for $n_1 = n_2$. For the summation of the correlation with double indices, the technique in [24, ch. 5] can be applied. Then, (67) becomes

$$\begin{aligned} \chi(i, j) &= \frac{1}{N-k} \sum_{t=-N+k+1}^{N-k-1} \left(1 - \frac{|t|}{N-k} \right) \\ &\cdot \gamma_{i,i}(tP) \gamma_{j,j}^*(tP) + \frac{1}{(N-k)^2} \\ &\cdot \sum_{n=1}^{N-k} \left[\gamma_{i,i}(0) E \left\{ \mathbf{a}^{*(j)}(n) \mathbf{a}^{(j)}(n) \right\} + \gamma_{j,j}^*(0) \right. \\ &\quad \cdot E \left\{ \mathbf{a}^{(j)}(n+k) \mathbf{a}^{*(i)}(n+k) \right\} \\ &\quad \left. + \frac{1}{(N-k)^2} \sum_{n=1}^{N-k} E \left\{ \mathbf{a}^{(i)}(n+k) \mathbf{a}^{*(i)}(n+k) \right\} \right. \\ &\quad \left. \cdot E \left\{ \mathbf{a}^{*(j)}(n) \mathbf{a}^{(j)}(n) \right\} \right]. \quad (68) \end{aligned}$$

By noticing (54), we can finally obtain (58). \square

REFERENCES

- [1] A. Aghamohammadi, H. Meyr, and G. Ascheid, "A new method for phase synchronization and automatic gain control of linearly modulated signals on frequency-flat fading channels," *IEEE Trans. Commun.*, vol. 39, pp. 25–29, Jan. 1991.
- [2] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.
- [3] J. Catipovic and J. G. Proakis, "Adaptive multichannel combining and equalization for underwater acoustic communications," *J. Amer. Statist. Assoc.*, vol. 94, pp. 1621–1631, Sept. 1993.
- [4] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 686–693, Nov. 1991.
- [5] —, "Pilot symbol assisted modulation and differential detection in fading and delay spread," *IEEE Trans. Commun.*, vol. 43, pp. 2206–2212, July 1995.
- [6] H. A. Cirpan and M. K. Tsatsanis, "Maximum likelihood blind channel estimation in the presence of Doppler shifts," *IEEE Trans. Signal Processing*, vol. 47, pp. 1559–1569, June 1999.
- [7] L. Davis, I. Collings, and R. Evans, "Identification of time-varying linear channels," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, vol. 5, Munich, Germany, Apr. 1997, pp. 3921–3924.
- [8] L. M. Davis, I. B. Collings, and R. J. Evans, "Constrained maximum likelihood estimation of time-varying linear channels," in *Proc. Signal Process. Adv. Wireless Commun.*, La Villette, France, April 16–18, 1997, pp. 1–4.
- [9] T. Eyceoz and A. Duel-Hallen, "Simplified block adaptive diversity equalizer for cellular mobile radio," *IEEE Commun. Lett.*, vol. 1, p. 15, Jan. 1997.
- [10] T. Eyceoz, A. Duel-Hallen, and H. Hallen, "Deterministic channel modeling and long range prediction of fast fading mobile radio channels," *IEEE Commun. Lett.*, vol. 2, p. 254, Sept. 1998.
- [11] —, "Prediction of fast fading parameters by resolving the interference pattern," in *Proc. 31st Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 2–5, 1997.
- [12] 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.
- [13] A. Goldsmith and P. P. Variaya, "Capacity of fading channels with channel side information," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1986–1992, Nov. 1997.
- [14] A. Goldsmith and S. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, pp. 1218–1230, Oct. 1997.
- [15] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 1996.

- [16] 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.
- [17] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.
- [18] M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Simulation of Communication Systems*. New York: Plenum, 1992.
- [19] S. M. Kay, *Modern Spectral Estimation: Theory and Application*. Englewood Cliffs, NJ: Prentice Hall, 1988.
- [20] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, 2nd ed. San Diego, CA: Academic, 1985.
- [21] C. Liu and K. Feher, "Pilot-symbol aided coherent M -ary PSK in frequency-selective fast Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 42, pp. 54–62, Jan. 1994.
- [22] H. Liu and G. B. Giannakis, "Deterministic approaches for blind equalization of time-varying channels with antenna arrays," *IEEE Trans. Signal Processing*, vol. 46, pp. 3003–3013, Nov. 1998.
- [23] A. J. Paulraj, C. Papadias, V. Reddy, and A. van der Veen, "Blind space-time signal processing," in *Wireless Communications: Signal Processing Perspectives*, H. V. Poor and G.W. Wornell, Eds. Englewood Cliffs, NJ: Prentice-Hall, 1998, ch. 4.
- [24] M. B. Priestley, *Spectral Analysis and Time Series*. San Diego, CA: Academic, 1981.
- [25] J. G. Proakis, *Digital Communications*, 2nd ed. New York: McGraw-Hill, 1989.
- [26] T. S. Rappaport, *Wireless Communications*. Upper Saddle River, NJ: Prentice-Hall, 1996.
- [27] S. Sampei and T. Sunaga, "Rayleigh fading compensation method for 16 QAM in digital land mobile radio channels," in *Proc. IEEE 39th Veh. Technol. Conf.*, San Francisco, CA, May 1989, pp. 640–646.
- [28] M. K. Tsatsanis, G. B. Giannakis, and G. Zhou, "Estimation and equalization of fading channels with random coefficients," *Signal Process.*, vol. 53, no. 2/3, Aug. 1996.
- [29] M. K. Tsatsanis and G. B. Giannakis, "Modeling and equalization of rapidly fading channels," *Int. J. Adaptive Contr. Signal Process.*, vol. 10, no. 2/3, pp. 159–176, Mar. 1996.
- [30] —, "Equalization of rapidly fading channels: Self-Recovering Methods," *IEEE Trans. Commun.*, vol. 44, pp. 619–630, May 1996.
- [31] A.-J. van der Veen, M. C. Vanderveen, and A. Paulraj, "Joint angle and delay estimation using shift-invariance techniques," *IEEE Trans. Signal Processing*, vol. 46, pp. 405–418, Feb. 1998.
- [32] J. Yang and K. Feher, "A digital Rayleigh fade compensation technology for coherent OQPSK system," in *Proc. IEEE 40th Veh. Technol. Conf.*, Orlando, FL, May 1990, pp. 732–737.



Michail K. Tsatsanis (M'92) received the diploma degree in electrical engineering from the National Technical University of Athens, Athens, Greece, in 1987 and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Virginia, Charlottesville, in 1990 and 1993, respectively.

He is currently an Associate Professor in the Electrical and Computer Engineering Department, Stevens Institute of Technology, Hoboken, NJ. His general research interests lie in the areas of statistical signal and array processing, with applications to communications and networking. His current interests focus on signal processing techniques for wireless communications including blind equalization, multiuser detection, fading channel estimation and tracking, and signal processing methods for networking problems.

Dr. Tsatsanis is a Member of the IEEE Technical Committee on SPCOM. He has served as a Member of the Organizing Committee for the 1996 IEEE Signal Processing Workshop on SSAP and as the Technical Co-Chair of the Organizing Committee for the 1999 IEEE Workshop on Signal Processing Advances in Wireless Communications. He received the 1998 NSF CAREER Award and the 1999 SP Society Young Author Best Paper Award. He is an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and IEEE COMMUNICATIONS LETTERS.



Zhengyuan (Daniel) Xu (M'99) received both the B.E. and M.E. degrees in electronic engineering from Tsinghua University, Beijing, China, in 1989 and 1991, respectively. He received the Ph.D. degree in electrical engineering from Stevens Institute of Technology, Hoboken, NJ, in 1999.

From 1991 to 1996, he worked as a System Engineer and Department Manager at the Tsinghua Unisplendour Group Corporation, Tsinghua University. From 1996 to 1999, he was a Research Assistant and Research Associate, respectively, with Stevens Institute of Technology, working on signal processing for wireless communications, especially for multiuser CDMA systems. Then, he joined the Department of Electrical Engineering, University of California, Riverside, as an Assistant Professor. His current research interests include digital communication theory, multiuser detection, system parameter estimation, and high-order statistics.

Dr. Xu received the Outstanding Student Award and Motorola Scholarship from Tsinghua University. He also received the Peskin Award from Stevens Institute for his outstanding performance.