Training-Based Time-Delay Estimation for CPM Signals Over Time-Selective Fading Channels

Ron Dabora, Jason Goldberg, *Senior Member, IEEE*, and Hagit Messer, *Fellow, IEEE*

***Abstract*—In this paper, we consider training-based symbol timing synchronization for continuous phase modulation over channels subject to flat, Rayleigh fading. A high signal-to-noise-ratio maximum-likelihood estimator based on a simplified channel correlation model is derived. The main objective is to reduce algorithm complexity to a single-dimensional search on the delay parameter, similar to that of the static-channel (slow fading) estimator. The asymptotic behavior of the algorithm is evaluated, and comparisons are made with the Cramer–Rao lower bound for the problem. Simulation results demonstrate highly improved performance over the conventional, static-channel delay estimator.**

***Index Terms*—Continuous phase modulation (CPM), fading channels, maximum-likelihood (ML) estimation, synchronization, timing.**

I. Introduction

CONTINUOUS phase modulation (CPM) is an important class of digital modulation that combines good spectral efficiency with the desirable property of constant signal modulus. This latter characteristic enables the use of highly efficient, nonlinear power amplification in transmission, and provides inherent robustness to amplitude fading in reception, e.g., [1]. Due to the popularity of CPM, considerable effort has been directed toward the problem of time synchronization for such signals, e.g., [2] and [3]. Nevertheless, most of the work on this subject has concentrated on the additive white Gaussian noise (AWGN) channel, with relatively little research directly addressing CPM time synchronization over *fast-fading* (i.e., time-varying) channels (see [4] for an exception).

Also, while the general problem of time synchronization over fast-fading channels has been considered in the literature, e.g., [2] and [5]–[7], some of this work either neglects the statistical nature of the fading or imposes additional signal assumptions. For example, in [2], a high signal-to-noise ratio (SNR) approximation removes the influence of the fading statistics from the problem, and in [5], the channel is simply treated as deterministically time varying, while in [7], a low-SNR-type "low energy coherence" assumption is used. Time synchronization for linear modulation types is considered in [8] and [9], where a cyclostationary approach is used, and in [10], which uses a delay-and-multiply method. In [11], a cyclostationary approach is applied to blind synchronization of minimum-shift keying (MSK) signals over time-selective fading channels. However, such methods require long data sequences, and are therefore not applicable to burst communications, for example.

The inherent limitations in time-delay estimation over flat, Rayleigh fading channels were investigated by several authors. In [9], the Cramer–Rao lower bound (CRLB) is derived for blind time-delay estimation for *linear* modulation by calculating the covariance of the transmitted signal (which is restricted to be real) and then imposing Gaussianity on the received signal. A general study of the bound on phase parameters for random amplitude phase-modulated signals can be found in [12] for a real random component, and in [13] for a complex random component. Finally, a detailed analysis of the inherent limitations in time synchronization of CPM for fast-fading channels can be found in [14]. In this paper, we present a simple estimation procedure derived through an approximate, high-SNR maximum-likelihood (ML) approach based on a simplified (mismatched) model for the channel fading process. The estimation procedure requires only a single-dimensional parameter search. Specifically, we focus on *training-based* synchronization, wherein the transmitted symbol sequence is *a priori* known at the receiver. Such a scenario may arise, for example, in communication systems which use known pilot transmission to synchronize the receiver to the transmitter. Another possible application is burst transmission, where a known preamble is periodically transmitted to enable correction of the receiver's reference clock.

Finally, we note that an important motivation for presenting this estimation procedure is that the simple, slow-fading estimator has considerably poorer performance than the proposed method, while both methods have approximately the same complexity. This is demonstrated in the following through simulations.

The rest of this paper is structured as follows. Section II reviews CPM, compares two fast-fading correlation models, and specifies the overall model of the received signal. The problem to be solved is then formulated. In Section III, the ML estimator is derived for a high SNR approximation, and in Section IV, simulation results are presented. Last, Section V summarizes the paper.

Paper approved by R. Reggiannini, the Editor for Synchronization and Wireless Applications of the IEEE Communications Society. Manuscript received June 17, 2002; revised May 27, 2003. This work was supported in part by the Israeli Science Foundation, founded by the Academy of Sciences and Humanities.

R. Dabora is with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853 USA (e-mail: rd97@cornell.edu).

J. Goldberg is with Phonetact, Inc., Los Altos, CA 94002 USA (e-mail: jason@eng.tau.ac.il)

H. Messer is with the Department of Electrical Engineering-Systems, Tel Aviv University, Tel Aviv 69978, Israel (e-mail: messer@eng.tau.ac.il).

Digital Object Identifier 10.1109/TCOMM.2004.831340

II. Modeling and Problem Formulation

Begin by considering a general, discrete-time complex baseband signal received over a time-selective fading channel and sampled at sampling interval T_s

$$y_i = f_i s_i(\tau) + n_i \tag{1}$$

where $f_i, s_i(\tau) = s(iT_s - \tau)$, and n_i are the baseband-equivalent time-varying channel, transmitted signal subject to unknown delay τ, and additive noise components, respectively.

The channel component is assumed to be a realization of a zero-mean, stationary, circular, complex Gaussian random process, $f_i \sim \mathcal{CN}(0, \sigma_f^2)$, with correlation function $R_f[m] \triangleq E\{f_i f_{i-m}^*\}$, where $(\cdot)^*$ denotes complex conjugation. Gaussianity of the channel gives rise to the well-known Rayleigh-distributed amplitude fading. The additive noise is assumed to be a realization of a sequence of independent, identically distributed (i.i.d.), zero-mean, circular, complex Gaussian random variables, $n_i \sim \mathcal{CN}(0, \sigma_n^2)$.

The continuous-time complex envelope of a CPM waveform may be written as, e.g., [15], $s(t) = \sqrt{(E_s/T_{\text{sym}})} e^{j[\varphi(t,\boldsymbol{\eta})+\varphi_0]}$, where E_s is the symbol energy, T_{sym} is the symbol duration, $\varphi(t, \boldsymbol{\eta})$ is the information-bearing phase, φ_0 is an arbitrary phase shift (which can be incorporated into the fading process and is thus taken to be zero), and $\boldsymbol{\eta} = [\ldots, \eta_{-2}, \eta_{-1}, \eta_0, \eta_1, \ldots]^T$ (where $[\cdot]^T$ denotes the vector transpose operation) is an infinitely long column vector of data symbols. Note that we use the convention $s_r(t) = \sqrt{2}\Re\{s(t)e^{j\omega_c t}\}$, where $s_r(t)$ and ω_c denote the real signal and the carrier frequency, respectively, and $\Re\{\cdot\}$ denotes the real part of the specified element.

A. Channel-Fading Statistics

A commonly used model for land-mobile communication scenarios with isotropic scattering and horizontal propagation is the "U-shaped" Jakes Doppler spectrum [16]

$$S_f^J(\omega) = \begin{cases} \sigma_f^2 \frac{1}{\omega_m/2} \frac{1}{\sqrt{1-\left(\frac{\omega}{\omega_m}\right)^2}}, & |\omega| \le \omega_m \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

where ω_m is the maximum Doppler shift in radians/second, which for the above model also corresponds to the Doppler bandwidth, B_d, through $\omega_m = 2\pi B_d$. The associated correlation sequence is

$$R_f^J(k) = \sigma_f^2 J_0(\omega_m k T_s) \tag{3}$$

where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind.

Though common, this representation of the fading process correlation was not found to enable derivation of a simple, ML-oriented estimate of the time delay. Thus, to facilitate the derivation of a simple estimator, we choose to model the channel as an autoregressive process (e.g., [17]) of order one (AR1).[1] Under the AR1 channel model, the channel correlation function may be written as

$$R_f^{\text{AR}}[m] = \sigma_f^2 \beta^{T_s|m|} = \sigma_f^2 \alpha^{|m|} \tag{4}$$

where β is the correlation parameter of the continuous-time fading process, and $\alpha = \beta^{T_s}$ can be viewed as an "effective correlation parameter" induced by the channel as well as the sampling rate. We define α of the AR1 model to be the normalized Jakes correlation at lag one

$$\alpha \triangleq J_0(\omega_m T_s). \tag{5}$$

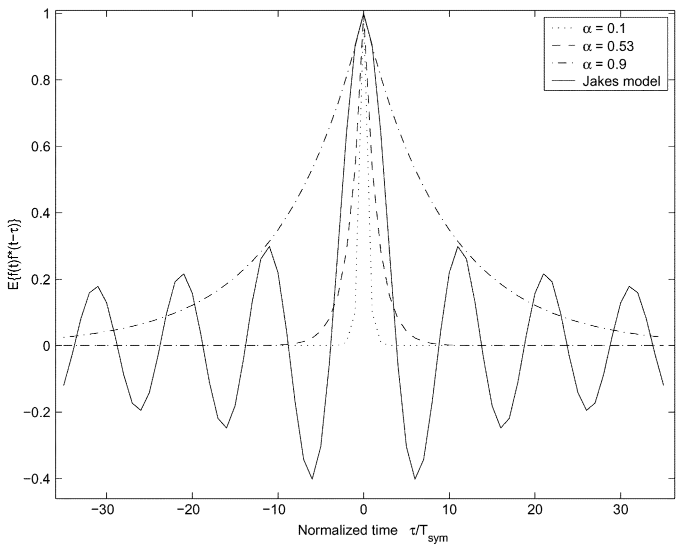

Fig. 1. Temporal correlation of the fading process for $B_d T_{\text{sym}} = 0.1$, for Jakes model and for AR1 fading correlation model, $\sigma_f^2 = 1, T_s = T_{\text{sym}}, \alpha = J_0(\omega_m T_s) = 0.9037$.

To aid in visualizing the difference between the two models, Fig. 1 depicts their respective temporal correlation functions. As can be seen from the figure, the correlation produced by the AR1 model lacks the oscillations present in the Jakes model, so there is no strong peak at ω_m and the spectrum decays more slowly with frequency. Use of the AR1 model when the channel is governed by (3) creates a type of model mismatch, the effect of which will be analyzed in the following.

B. Statistics of the Received Signal

Assume an observation interval of N samples giving rise to $\{y_i\}_{i=0}^{N-1}$ which, from (1), may be written in matrix form as

$$\begin{aligned} \mathbf{y} &= [y_0, \ldots, y_{N-1}]^T = \mathbf{Sf} + \mathbf{n} \\ \mathbf{f} &= [f_0, \ldots, f_{N-1}]^T \\ \mathbf{S} &= \text{diag}[\mathbf{s}], \quad \mathbf{s} = [s_0(\tau), s_1(\tau), \ldots, s_{N-1}(\tau)]^T \\ \mathbf{n} &= [n_0, \ldots, n_{N-1}]^T \end{aligned} \tag{6}$$

with $\text{diag}[\cdot]$ denoting a diagonal matrix of specified diagonal elements. Before continuing with the derivation, note that since it is impossible to distinguish between the fading channel power σ_f^2 and the signal power E_s/T_{sym}, we will use σ_f^2 to represent their product (i.e., the signal power is normalized to one), such that the SNR is defined as $\rho = \sigma_f^2/\sigma_n^2$. Since the transmitted signal $\mathbf{s}$ is known up to the fixed time delay, $\mathbf{y}$ is a zero-mean complex Gaussian random vector, with correlation matrix given by

$$\begin{aligned} E\{\mathbf{yy}^H\} = \mathbf{R_y} &= \mathbf{SR_f S}^H + \sigma_n^2 \mathbf{I} \\ \mathbf{R_f} &= E\{\mathbf{ff}^H\} \end{aligned} \tag{7}$$

where $[\cdot]^H$ denotes conjugate transposition, $\mathbf{I}$ stands for the $N \times N$ identity matrix, and $\mathbf{R_f}$ is the fading-channel correlation matrix generated according to the Jakes model. The *true* (i.e., not

[1]Note that, at least for frequency estimation, both [17] and [18] claim that the actual shape of the Doppler spectrum has no noticeable effect on performance.

mismatched) probability density function (pdf) of the received signal can now be written as

$$p(\mathbf{y};\boldsymbol{\theta}_t) = \frac{1}{\pi^N |\mathbf{R_f} + \sigma_n^2 \mathbf{I}|} e^{-\mathbf{y}^H \mathbf{S}(\mathbf{R_f}+\sigma_n^2\mathbf{I})^{-1}\mathbf{S}^H \mathbf{y}}$$
$$\boldsymbol{\theta}_t = \left[\tau, \sigma_f^2, \sigma_n^2, \omega_m\right]^T \tag{8}$$

where $|\cdot|$ denotes the determinant of a matrix, and we used the fact that for CPM, the signal matrix $\mathbf{S}$ is unitary.

C. Performance Bound

In Section IV, the performance of the proposed algorithm is compared with the performance of the conventional slow-fading delay-estimation algorithm and with the CRLB evaluated for both the Jakes model and the AR1 channel approximation. The CRLB is a bound commonly used for the performance evaluation of unbiased estimation algorithms, and is evaluated as follows. Consider a data vector $\mathbf{y}$ whose pdf $p(\mathbf{y};\boldsymbol{\theta}_t)$ is parameterized by a parameter vector $\boldsymbol{\theta}_t$. Under a set of regularity conditions, the CRLB for the mean square error (MSE) of any unbiased estimator $\hat{\boldsymbol{\theta}}_{t,i}(\mathbf{y})$, of the ith parameter $\boldsymbol{\theta}_{t,i}$ is given by, e.g., [19]

$$E\{(\hat{\boldsymbol{\theta}}_{t,i}(\mathbf{y}) - \boldsymbol{\theta}_{t,i})^2\} \geq [\mathbf{J}^{-1}(\boldsymbol{\theta}_t)]_{ii} \tag{9}$$

where the matrix $\mathbf{J}(\boldsymbol{\theta}_t)$ is known as the Fisher information matrix (FIM), whose elements are given by

$$[\mathbf{J}(\boldsymbol{\theta}_t)]_{ij} = E\left\{\frac{\partial p(\mathbf{y};\boldsymbol{\theta}_t)}{\partial \boldsymbol{\theta}_{t,i}} \left(\frac{\partial p(\mathbf{y};\boldsymbol{\theta}_t)}{\partial \boldsymbol{\theta}_{t,j}}\right)^*\right\} \tag{10}$$

where $[\mathbf{A}]_{ij}$ denotes the (i,j)th element of the matrix $\mathbf{A}$.

D. Estimation Problem

The estimation problem can now be formulated as follows. Given a received signal $\mathbf{y}$ with pdf (8), which is parameterized by an unknown parameter vector, $\boldsymbol{\theta}_t$, estimate the time delay, τ. In this problem, τ is the parameter of interest, and the three remaining parameters are *nuisance* parameters.

III. ML Estimator

A. AR1 Approximation of the Likelihood Function

For the AR1 model, use of (4) yields the following correlation matrix:

$$\mathbf{R_f} = \sigma_f^2 \begin{bmatrix} 1 & \alpha & \alpha^2 & \cdots & \alpha^{N-1} \\ \alpha & 1 & \alpha & \cdots & \alpha^{N-2} \\ \alpha^2 & \alpha & 1 & \cdots & \alpha^{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha^{N-1} & \alpha^{N-2} & \alpha^{N-3} & \cdots & 1 \end{bmatrix}. \tag{11}$$

The pdf of (8) with (11) replacing the Jakes correlation matrix is the pdf assumed by the estimator. The "mismatched" parameter vector is now defined as

$$\boldsymbol{\theta} = \left[\tau, \boldsymbol{\theta}_n^T\right]^T, \qquad \boldsymbol{\theta}_n = \left[\sigma_f^2, \sigma_n^2, \alpha\right]^T. \tag{12}$$

The relation between α and ω_m given by (5) is now seen to be a natural consequence of the AR1 approximation for the channel correlation used by the mismatched estimator.

Before describing the estimation procedure, we wish to make a few remarks on the estimation of ω_m. Although in this paper, we are interested only in delay estimation and not in estimating the Doppler bandwidth per se, as will be shortly explained, its estimation can be obtained from the intermediate steps of the delay-estimation procedure. Due to the introduction of the AR1 correlation model, our estimation procedure will produce the value of $J_0(\omega_m T_s)$. This value is enough for estimating the delay, but if we would like to estimate the Doppler bandwidth, we must use the inverse Bessel function. However, the inverse Bessel function is not a one-to-one mapping. It would be a one-to-one mapping only if we limit ourselves to the region from the first maxima (at $\omega_m T_s = 0$) to the first minima (at $\omega_m T_s = 3.83$). Hence, using $\omega_m T_s = 2\pi B_d T_{\text{sym}}/K_s$, with $K_s = (T_{\text{sym}}/T_s)$, we conclude that for $B_d T_{\text{sym}} \leq 0.6$ (up to the first minimum of the Jakes correlation function) and $T_s \leq T_{\text{sym}}, \alpha$ and ω_m have a one-to-one relation in the sense that given one of these parameters, the other can be uniquely specified. Thus, for virtually all practical scenarios, searching the $\boldsymbol{\theta}$ parameter domain is "equivalent" to searching the $\boldsymbol{\theta}_t$ parameter domain.

B. Outline of Approach

The straightforward approach for deriving the ML solution to the problem is to try to compress (8) (where (11) is used instead of the true channel correlation matrix) with respect to the nuisance parameters, and to perform a search on the delay and the parameters which cannot be compressed. This approach has been taken in [20], where using high and low SNR approximations, we were able to reduce the search to a two-dimensional search in τ and α. Also, note that [21], which considers the problem of estimating the direction of arrival (DOA) in the presence of correlated multiplicative AR1 spatial noise, derives a double search in their "equivalent" τ and α for low SNR, while for high SNR, an explicit solution for the ML estimate is derived. However, due to the particularly simple structure of the $\mathbf{S}$ matrix in the DOA problem of [21], such a solution is not applicable to the current problem. Thus, we turned to another approach: applying the chain rule for expressing a pdf in terms of the product of conditional pdfs, we show that it is possible to reduce the problem, under a high SNR approximation, to a *single-parameter search*.

Mathematically, the compression can be expressed as

$$\hat{\tau} = \arg\max_{\tau}\{\bar{\Lambda}(\tau;\mathbf{y})\} \tag{13}$$
$$\bar{\Lambda}(\tau;\mathbf{y}) = \max_{\boldsymbol{\theta}_n}\{\Lambda(\boldsymbol{\theta}_n,\tau;\mathbf{y})\}$$
$$\Lambda(\boldsymbol{\theta}_n,\tau;\mathbf{y}) = W_0 - \log\left|\mathbf{R_f} + \sigma_n^2\mathbf{I}\right| - \mathbf{y}^H\mathbf{S}\left(\mathbf{R_f} + \sigma_n^2\mathbf{I}\right)^{-1}\mathbf{S}^H\mathbf{y}$$

where $\Lambda(\boldsymbol{\theta}_n,\tau;\mathbf{y})$ is the log-likelihood function for the problem, $W_0 = -N\log(\pi)$, and $\mathbf{R_f}$ is defined in (11). The next subsections concentrate on simplified computation of $\bar{\Lambda}(\tau;\mathbf{y})$.

C. Evaluating the Conditional PDF

Our objective is to compress the estimation problem into a one-dimensional search on the delay. Thus, the nuisance parameters are to be expressed in terms of the received signal $\mathbf{y}$ and the delay τ. Since the received symbols are *a priori* known, by fixing τ we can write $\mathbf{z} \triangleq \mathbf{S}^H\mathbf{y}$, which for a CPM signal, is a unitary transformation of $\mathbf{y}$. This implies that

$$\begin{aligned}\bar{\Lambda}(\tau;\mathbf{y}) &= \max_{\boldsymbol{\theta}_n}\{\Lambda(\boldsymbol{\theta}_n,\tau;\mathbf{z}=\mathbf{S}^H\mathbf{y})\}\\ &\triangleq \max_{\boldsymbol{\theta}_n}\{\bar{\Lambda}(\boldsymbol{\theta}_n;\mathbf{z})\}.\end{aligned} \tag{14}$$

We seek to find the solution $\hat{\boldsymbol{\theta}}_n = \arg\max_{\boldsymbol{\theta}_n}\{\bar{\Lambda}(\boldsymbol{\theta}_n;\mathbf{z})\}$.

Applying the transformation $\mathbf{z} \triangleq \mathbf{S}^H\mathbf{y}$ to (8) yields an expression for $\bar{\Lambda}(\boldsymbol{\theta}_n;\mathbf{z})$ of the form

$$\bar{\Lambda}(\boldsymbol{\theta}_n;\mathbf{z}) = W_0 - \log\left|\mathbf{R_f}+\sigma_n^2\mathbf{I}\right| - \mathbf{z}^H\left(\mathbf{R_f}+\sigma_n^2\mathbf{I}\right)^{-1}\mathbf{z}.$$

The AR1 model allows us to use the chain rule, and limit the "history dependence" of z_i to z_{i-1} only. Derivation of the (Gaussian) likelihood functions $\tilde{p}(z_0;\boldsymbol{\theta}_n)$ and $\tilde{p}(z_i|z_{i-1};\boldsymbol{\theta}_n)$ yields (see Appendix A)

$$\begin{aligned}\bar{\Lambda}(\boldsymbol{\theta}_n;\mathbf{z}) &= \log(\tilde{p}(z_0;\boldsymbol{\theta}_n))\\ &\quad+\sum_{i=1}^{N-1}\log(\tilde{p}(z_i|z_{i-1};\boldsymbol{\theta}_n))\end{aligned} \tag{15}$$

$$\tilde{p}(z_0;\boldsymbol{\theta}_n) = \frac{1}{\pi\left(\sigma_f^2+\sigma_n^2\right)}e^{-\frac{|z_0|^2}{\sigma_f^2+\sigma_n^2}} \tag{16}$$

$$\tilde{p}(z_i|z_{i-1};\boldsymbol{\theta}_n) = \frac{1}{\pi\left(\sigma_f^2+\sigma_n^2-\frac{(\alpha\sigma_f^2)^2}{\sigma_f^2+\sigma_n^2}\right)}e^{-\frac{\left|z_i-z_{i-1}\frac{\alpha\sigma_f^2}{\sigma_f^2+\sigma_n^2}\right|^2}{\sigma_f^2+\sigma_n^2-\frac{(\alpha\sigma_f^2)^2}{\sigma_f^2+\sigma_n^2}}}.$$

The complete likelihood function for the estimation of the nuisance parameters can now be written as

$$\begin{aligned}\bar{\Lambda}(\boldsymbol{\theta}_n;\mathbf{z}) &= W_0 - \log\left(\sigma_n^2+\sigma_f^2\right) - \frac{|z_0|^2}{\sigma_f^2+\sigma_n^2}\\ &\quad+\sum_{i=1}^{N-1}\left\{-\log\left(\sigma_f^2+\sigma_n^2-\frac{\left(\alpha\sigma_f^2\right)^2}{\sigma_f^2+\sigma_n^2}\right)\right.\\ &\quad\left.-\frac{\left|z_i-z_{i-1}\frac{\alpha\sigma_f^2}{\sigma_f^2+\sigma_n^2}\right|^2}{\sigma_f^2+\sigma_n^2-\frac{\left(\alpha\sigma_f^2\right)^2}{\sigma_f^2+\sigma_n^2}}\right\}.\end{aligned} \tag{17}$$

Differentiating (17) with respect to α and equating to zero results in a third-order polynomial for the estimation of α, which depends on the other two nuisance parameters, σ_f^2 and σ_n^2, and on the delay and the received signal through $\mathbf{z}$. This requires a simultaneous solution of the likelihood equations for the nuisance parameters σ_f^2 and σ_n^2 at any given τ. However, since we are interested in a simpler estimation procedure, we resort to a high-SNR approximation.

D. High-SNR Approximation

For the high-SNR approximation, we assume that $\sigma_n^2+\sigma_f^2 \approx \sigma_f^2$. The high-SNR approximation allows us to omit σ_n^2 from the nuisance parameters vector, leaving us with $\boldsymbol{\theta}_n^{\text{high}} = [\sigma_f^2,\alpha]^T$. The likelihood function of (17) under this approximation becomes

$$\begin{aligned}\tilde{\Lambda}^{\text{high}}\left(\boldsymbol{\theta}_n^{\text{high}};\mathbf{z}\right) &\approx W_0 - \log\left(\sigma_f^2\right) - \frac{|z_0|^2}{\sigma_f^2}\\ &\quad-\sum_{i=1}^{N-1}\left\{\log\left(\sigma_f^2(1-\alpha^2)\right)\right.\\ &\quad\left.+\frac{|z_i-\alpha z_{i-1}|^2}{\sigma_f^2(1-\alpha^2)}\right\}\end{aligned} \tag{18}$$

where it is assumed that $|\alpha| \neq 1$.

Straightforward differentiation of (18) can now be used to obtain closed-form estimators for σ_f^2 and α. Begin with σ_f^2

$$\begin{aligned}&\frac{\partial\tilde{\Lambda}^{\text{high}}\left(\boldsymbol{\theta}_n^{\text{high}};\mathbf{z}\right)}{\partial\sigma_f^2}\\ &\quad=\sum_{i=1}^{N-1}\left\{-\frac{1}{\sigma_f^2}+\frac{1}{\left(\sigma_f^2\right)^2}\frac{|z_i-\alpha z_{i-1}|^2}{1-\alpha^2}\right\}\\ &\quad\quad-\frac{1}{\sigma_f^2}+\frac{|z_0|^2}{\left(\sigma_f^2\right)^2}=0.\end{aligned} \tag{19}$$

Extracting σ_f^2 yields

$$\begin{aligned}\hat{\sigma}_f^2 = \frac{1}{N}\Bigg[|z_0|^2+\frac{1}{1-\alpha^2}\sum_{i=1}^{N-1}&(|z_i|^2+\alpha^2|z_{i-1}|^2\\ &-\alpha(z_iz_{i-1}^*+z_i^*z_{i-1}))\Bigg].\end{aligned} \tag{20}$$

In (20), $\hat{\sigma}_f^2$ is expressed in terms of α and $\mathbf{z}$. We now calculate the high-SNR ML equation for α by first writing

$$\begin{aligned}&\frac{\partial\tilde{\Lambda}^{\text{high}}\left(\boldsymbol{\theta}_n^{\text{high}};\mathbf{z}\right)}{\partial\alpha}\\ &=\sum_{i=1}^{N-1}\left\{-\frac{1}{1-\alpha^2}(-2\alpha)+\frac{|z_i-\alpha z_{i-1}|^2}{\sigma_f^2(1-\alpha^2)^2}(-2\alpha)\right.\\ &\quad\left.-\frac{(z_i-\alpha z_{i-1})(-z_{i-1}^*)}{\sigma_f^2(1-\alpha^2)}-\frac{(z_i^*-\alpha z_{i-1}^*)(-z_{i-1})}{\sigma_f^2(1-\alpha^2)}\right\}=0.\end{aligned} \tag{21}$$

Collecting the coefficients for each order of α we find that $\hat{\alpha}$ is given as the solution of:

$$-2\alpha^3\sigma_f^2(N-1)+\alpha^2B+\alpha\left(2\sigma_f^2(N-1)-2A\right)+B=0,$$

where $B = \sum_{i=1}^{N-1}(z_iz_{i-1}^*+z_{i-1}z_i^*)$ and $A = \sum_{i=1}^{N-1}(|z_{i-1}|^2+|z_i|^2)$. Observe that the estimation of α requires solving a third-order polynomial for $\hat{\alpha}$ that depends on σ_f^2 and $\mathbf{z}$. We can now plug in the expression for $\hat{\sigma}_f^2$ of (20) into this equation, yielding an equation for α only. Assume that N is large enough so that $\sigma_f^2(N-1) \approx \sigma_f^2 N$. From (20)

$$\sigma_f^2 N = D + \frac{1}{1-\alpha^2}(\alpha^2 A - \alpha B) \quad (22)$$

where $D \triangleq \sum_{i=0}^{N-1} |z_i|^2$. Hence, we have

$$\begin{aligned} &-2\frac{\alpha^3}{1-\alpha^2}((1-\alpha^2)D + \alpha^2 A - \alpha B) + \alpha^2 B \\ &+ 2\frac{\alpha}{1-\alpha^2}((1-\alpha^2)D + \alpha^2 A - \alpha B \\ &- (1-\alpha^2)A) + B = 0. \end{aligned}$$

Collecting coefficients of the same order of α yields

$$\begin{aligned} \alpha^5(2D-2A) + \alpha^4 B + \alpha^3(4A-4D) - 2\alpha^2 B \\ + \alpha(2D-2A) + B = 0. \end{aligned} \quad (23)$$

Define next

$$E = D - A = -\sum_{i=2}^{N-1} |z_{i-1}|^2 \quad (24)$$

so the ML equation for α can be written as

$$\begin{aligned} \alpha^5 2E + \alpha^4 B - \alpha^3 4E - 2\alpha^2 B + \alpha 2E + B = 0 \\ \Rightarrow (\alpha 2E + B)(\alpha^2 - 1)^2 = 0 \end{aligned} \quad (25)$$

resulting in an explicit high-SNR estimate for α

$$\hat{\alpha} = \frac{-B}{2E}. \quad (26)$$

Using (26) in (20), we obtain the estimate of σ_f^2 in terms of the input sequence and the delay as

$$\hat{\sigma}_f^2 = \frac{1}{N}\frac{1}{4E^2 - B^2}(4E^2 D + B^2 E). \quad (27)$$

Note that if f_i has a correlation following the Jakes model, then, at the true value of τ, the estimator for $\hat{\alpha}$ yields (recall that $\sigma_f^2 \gg \sigma_n^2$)

$$\begin{aligned} \hat{\alpha} = \frac{-B}{2E} &= \frac{-\sum_{i=1}^{N-1}(z_i z_{i-1}^* + z_{i-1} z_i^*)}{-2\sum_{i=2}^{N-1} |z_{i-1}|^2} \\ &\overset{N\to\infty}{\to} \frac{2(N-1)\sigma_f^2 J_0(\omega_m T_s)}{2(N-2)(\sigma_f^2 + \sigma_n^2)} \approx J_0(\omega_m T_s) \end{aligned} \quad (28)$$

further motivating the definition of α in (5).

Similarly, we can find the limiting estimate of σ_f^2 that the estimator produces. From (20), using (28) and the high-SNR condition $\sigma_f^2 \gg \sigma_n^2$, we get that

$$\begin{aligned} \hat{\sigma}_f^2 \overset{N\to\infty}{\to} &\frac{1}{N}\Bigg[\left(\sigma_f^2 + \sigma_n^2\right) + \frac{N-1}{1-\hat{\alpha}^2}\left(\sigma_f^2 + \sigma_n^2\right)(1+\hat{\alpha}^2) \\ &-2\hat{\alpha}^2\frac{N-2}{1-\hat{\alpha}^2}\left(\sigma_f^2 + \sigma_n^2\right)\Bigg] \\ \approx &\frac{1}{1-\hat{\alpha}^2}\left(\sigma_f^2 + \sigma_n^2\right)(1+\hat{\alpha}^2 - 2\hat{\alpha}^2) \\ = &\sigma_f^2 + \sigma_n^2 \approx \sigma_f^2. \end{aligned} \quad (29)$$

Note that (29) is not valid for $|\hat{\alpha}| = 1$.

The overall estimation procedure is now summarized as follows. For each value of τ in the search domain, we evaluate $\mathbf{z} = \mathbf{S}^H\mathbf{y}$. Next, use $\mathbf{z}$ in (26) and (27) to obtain the estimates of $\hat{\alpha}$ and $\hat{\sigma}_f^2$, respectively. The estimates of the nuisance parameters are then plugged into (18), and the log-likelihood value is evaluated for the corresponding τ. The delay estimate is that value of τ which maximizes $\tilde{\Lambda}^{\text{high}}(\boldsymbol{\theta}_n^{\text{high}}; \mathbf{z})$.

Thus, for high SNR, the nuisance parameters are given in closed-form expressions that depend solely on the delay and the received sequence, reducing the search to a single-parameter search on τ only.

The estimator derived above was derived under the assumption that $\alpha \neq 1$ ($\alpha = 1$ implies $\omega_m = 0$, the slow-fading case). Thus, the algorithm would produce valid results as long as $|\hat{\alpha}| \neq 1$. In that context, the actual value of α is not relevant. Whenever $|\hat{\alpha}| = 1$, the estimator will not produce a valid result. However, the estimate $\hat{\alpha}$, calculated using (26), is a random variable, thus $\Pr(|\hat{\alpha}| = 1) = 0$.[2] This implies that even for the case of $\alpha = 1$, the probability that the estimated channel correlation parameter is exactly one is zero. Hence, even when the value of the (*a priori* unknown) channel correlation parameter is one, we are able to apply the method above as an *ad-hoc* procedure. In the next section, we examine its performance for this case via simulation.

IV. Numerical Example

This section presents numerical examples which illustrate the performance of the estimator derived in this paper as a function of SNR, Doppler-time product, and sample size. A nominal MSK modulation scenario is considered with a bit rate of $R_{\text{sym}} = (1/T_{\text{sym}}) = 24$ Kb/s, a sampling rate of $T_s = (T_{\text{sym}}/2)$, an SNR of $\rho = 10$ dB, and a bit sequence of length $L_s = 41$ bits, alternating in pairs. This sequence is shown in [14] to minimize the CRLB for the problem. The Jakes channel correlation model with Doppler-time product of $B_d T_{\text{sym}} = 0.02$ (corresponding, via (5), to an AR1 fading channel model with $\alpha = 0.999$) is used to generate the channel. We note that the search in the simulations was restricted to delays in the range $[-2T_{\text{sym}}, 2T_{\text{sym}})$, since this is the length of the basic period of the MSK signal, hence, delays outside this interval will fold inside. The results are plotted as one of the above-listed parameters is varied from its nominal value. Each figure depicts the simulation results for the AR1 estimator, together with the CRLBs for the AR1 and the Jakes channel correlation models (see [14]), and an analytically computed asymptotic error variance (see Appendix B). The performance of the slow-fading matched-filter-type estimator, which assumes that the channel is a realization of an unknown complex Gaussian random variable, is also presented for comparison (see [6]).

Begin with Fig. 2, which presents the results as a function of the sequence length in symbols. We observe that the CRLB for the Jakes model is very close to that of the AR1 model. Also note that the asymptotic variance of the mismatched estimator is very close to the CRLB for the Jakes model. Therefore, we gain simplified estimation without a significant penalty in excess error. The variance of the AR1 estimator is seen to asymptotically approach the CRLB as the sequence length increases, and for 300 symbols sequence, the estimated variance is only 1.12 dB from the CRLB and only 0.82 dB from the analytic prediction. On the

[2]Unless the pdf of $\hat{\alpha}$ contains impulses at $+1$ or -1, which does not seem to be the case.

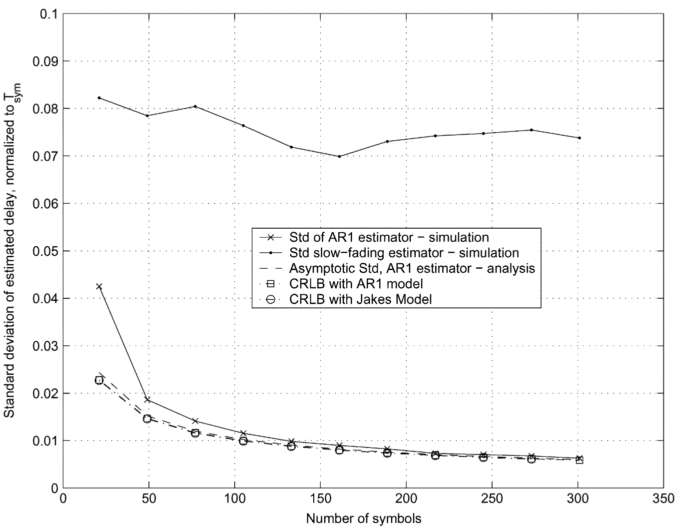

Fig. 2. Standard deviation of estimated delay versus sequence length, for Jakes channel, $R_{\rm sym} = (1/T_{\rm sym}) = 24$ Kb/s, $T_s = (T_{\rm sym}/2)$, SNR $= 10$ dB, $B_d T_{\rm sym} = 0.02$.

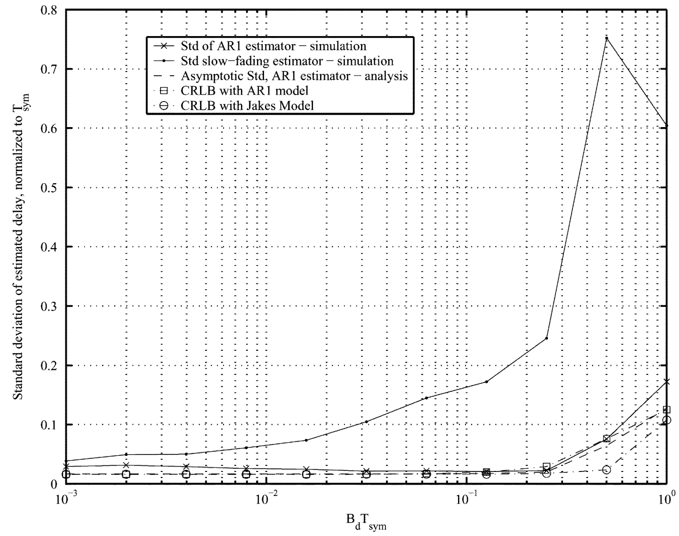

Fig. 4. Standard deviation of estimated delay versus Doppler-time product, for Jakes channel, $R_{\rm sym} = (1/T_{\rm sym}) = 24$ Kb/s, $T_s = (T_{\rm sym}/2)$, SNR $=$ 10 dB, $L_s = 41$ symbols.

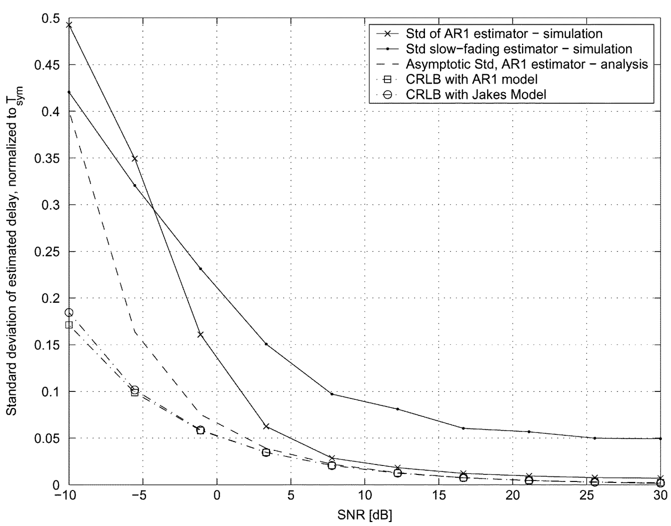

Fig. 3. Standard deviation of estimated delay versus SNR, for Jakes channel, $R_{\rm sym} = (1/T_{\rm sym}) = 24$ Kb/s, $T_s = (T_{\rm sym}/2)$, $B_d T_{\rm sym} = 0.02$, $L_s = 41$ symbols.

other hand, the slow-fading estimator is much worse—its standard deviation is 6–7 times that of the AR1 estimator.

Next, Fig. 3 presents the simulation results as the SNR varies from -10 to 30 dB. Examining the figure, we can see that both bounds are again quite close. We can also see that the asymptotic variance of the mismatched AR1 estimator is higher than the Jakes bound at low SNR, but converges to the Jakes bound at high SNR. Observe that the simulation results are quite close to the CRLB and agree well with the asymptotic analysis; the differences are because of the relatively small data-record length used (i.e., not "sufficiently asymptotic" conditions). Also note that the slow-fading estimator is considerably worse than the mismatched estimator for all practical SNRs. For very low SNRs (below -4 dB), however, we see that the slow-fading estimator has smaller variance than the AR1 estimator. This can be explained as follows. The estimator's performance is influenced by two factors, the fading process and the thermal noise. As long as the fading process is the dominating error factor, the AR1 estimator (which uses a high-SNR approximation) is superior to the slow-fading estimator, since it is better matched to the fading channel. However, at very low SNRs, the dominant error factor is the thermal noise. Here, the fact that the AR1 estimator assumes high SNR while the slow-fading estimator does not make any SNR approximation, plays in favor of the slow-fading estimator and enables it to perform better than the AR1 estimator. However, for the practical range of SNRs (say, above 0 dB), the AR1 estimator is superior to the slow-fading estimator.

Fig. 4 presents the simulation results as a function of the Doppler-time product at the nominal 10 dB SNR. We observe that as the Doppler-time product increases, the variance of the AR1 delay estimate remains quite constant up to Doppler-time product values of 0.2, which includes almost all practical scenarios. The peak of the slow-fading estimator around $B_d T_{\rm sym} = 0.5$ can be explained by examining the sampled Jakes correlation. Since this correlation function has zero crossings, there are combinations of the Doppler-time product and the sampling interval T_s that result in a sampled correlation with smaller values at nonzero lags, and other combinations that result in larger values of the sampled correlation at nonzero lags. For our specific scenario, the minimal correlation values are achieved when the Doppler-time product is near 0.5, and then, with the increase of the Doppler-time product, the correlation increases (and becomes negative at the first lag, due to the zero crossing of the Jakes correlation function). This implies that the estimator observes the least correlated fading process around $B_d T_{\rm sym} = 0.5$, and this is the reason for the peak in the variance. This is true also for the AR1 estimator, only there, this phenomenon becomes noticeable when the error is dominated by the fading process rather than the thermal noise, which is not the case presented in the figure.

Also, note the slight deviation from the asymptotic variance at very low Doppler-time product values. To examine the dependence of this behavior on SNR, the simulation was repeated at 20 dB SNR, as shown in Fig. 5. In Fig. 5, the dependence of estimation performance on the Doppler-time product values

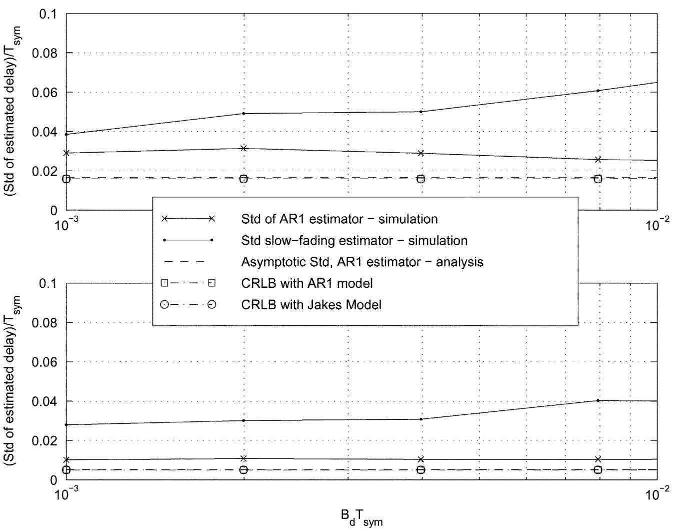

Fig. 5. Standard deviation of estimated delay versus Doppler-time product, for Jakes channel, $R_{\rm sym} = (1/T_{\rm sym}) = 24$ Kb/s, $T_s = (T_{\rm sym}/2), L_s = 41$ symbols, SNR = 10 dB (top), SNR = 20 dB (bottom).

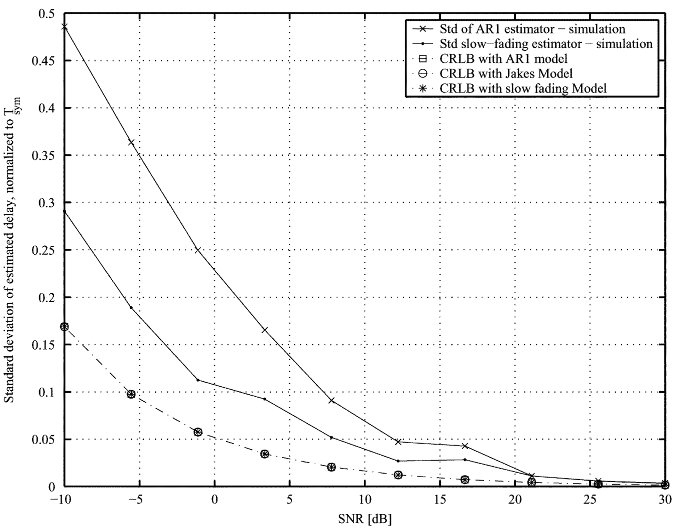

Fig. 6. Standard deviation of estimated delay versus SNR, for slow-fading channel, $R_{\rm sym} = (1/T_{\rm sym}) = 24$ Kb/s, $T_s = (T_{\rm sym}/2), B_d T_{\rm sym} = 0$, $L_s = 41$ symbols.

is presented for SNR of 10 dB at the top half and for SNR of 20 dB at the bottom half, zooming in on the interesting region of low Doppler-time product values. We observe that at SNR of 20 dB, the performance remains constant over a wide range of small Doppler-time product values. The reason for the deviation from the prediction in Fig. 4 can now be understood from (28). As the SNR decreases, the estimation of α becomes more and more biased. This effect, in turn, shifts the delay estimate from its true location, resulting in worse estimator performance. Furthermore, as $\alpha \rightarrow 1$, the likelihood cost, as a function of α, becomes steeper (see (21), where the term $1 - \alpha^2$ appears in the denominator), thus a small error in the value of α causes a relatively large change in the cost function. This sensitivity induces a large error variance in the estimate. Hence, for very low values of Doppler-time product, higher SNRs are required for the asymptotic performance to be attained.

Last, Fig. 6 investigates the performance as a function of SNR for the slow-fading (i.e., static) channel. As expected, the conventional static-channel delay estimator is optimal in the sense that its performance asymptotically achieves the bound. In the current scenario, the sequence length is too short for the asymptotic conditions to exist at low and moderate SNRs, however, the performance of the static-channel estimator is still the best. Nevertheless, the newly proposed estimator is seen to offer nearly optimal performance for the static channel, and for SNRs higher than 15 dB, its performance practically converges to the slow-channel estimator performance where both converge to the bound. It may therefore be applied in scenarios where it is not known *a priori* whether the channel fading is of time-varying nature.

V. Conclusion

This paper presented an ML approach for estimating time delay for CPM over Rayleigh flat-fading channels. The approach was based on a mismatched AR1 channel-correlation model upon which a high SNR estimator was derived. The estimator was compressed into a single-parameter search over the delay only. Numerical evaluation showed that the error variance is very close to the CRLB evaluated with the Jakes channel-correlation model. While both the newly proposed estimator and the conventional, static-channel estimator require a one-dimensional search over the delay parameter, simulation results clearly indicate the superior performance of the former over the latter.

Appendix A
Derivation of the Conditional PDF $\tilde{p}(z_i|z_{i-1};\boldsymbol{\theta}_n)$

Under the AR1 model, the received signal is described by the following equations:

$$\begin{aligned} y_i &= s_i f_i + n_i \\ f_i &= \alpha f_{i-1} + \sqrt{1-\alpha^2} w_i \end{aligned} \tag{A.1}$$

where, $w_i \sim \mathcal{CN}(0, \sigma_f^2)$, i.i.d. By fixing τ, the elements of the transformed vector $\mathbf{z}$ are given by

$$\begin{aligned} z_i \triangleq y_i s_i^* &= f_i + s_i^* n_i \\ &= \alpha f_{i-1} + \sqrt{1-\alpha^2} w_i + s_i^* n_i. \end{aligned} \tag{A.2}$$

Using

$$z_{i-1} = f_{i-1} + s_{i-1}^* n_{i-1} \Rightarrow f_{i-1} = z_{i-1} - s_{i-1}^* n_{i-1}$$

we can write

$$\begin{aligned} z_i &= \alpha z_{i-1} + \sqrt{1-\alpha^2} w_i + s_i^* n_i - \alpha s_{i-1}^* n_{i-1} \\ &= \alpha z_{i-1} + \sqrt{1-\alpha^2} w_i + \tilde{n}_i - \alpha \tilde{n}_{i-1} \end{aligned} \tag{A.3}$$

where

$$\tilde{n}_i \triangleq s_i^* n_i \sim \mathcal{CN}\left(0, \sigma_n^2\right). \tag{A.4}$$

Since $y_i;\boldsymbol{\theta}_t \sim \mathcal{CN}(0, \sigma_f^2 + \sigma_n^2)$, and z_i is a deterministic unitary transformation of y_i, then z_i is also distributed as complex normal: $z_i;\boldsymbol{\theta}_n \sim \mathcal{CN}(0, \sigma_f^2 + \sigma_n^2)$. Consider next the conditional pdf of z_i, namely $\tilde{p}(z_i|\mathbf{z}^{i-1};\boldsymbol{\theta}_n), \mathbf{z}^{i-1} = [z_0, z_1, \ldots, z_{i-1}]^T$. From (A.3), it is evident that the dependence of z_i on its history

$\mathbf{z}^{i-1}$, for the AR1 fading model, is limited to dependence on the previous sample z_{i-1} alone

$$\tilde{p}(z_i|\mathbf{z}^{i-1};\boldsymbol{\theta}_n) = \tilde{p}(z_i|z_{i-1};\boldsymbol{\theta}_n).$$

By definition, n_i and w_i are normal and independent of z_{i-1}; also, $z_{i-1}|z_{i-1}$ is a constant. Hence, from (A.3), it is evident that the only element required to be determined in order to evaluate the conditional distribution $\tilde{p}(z_i|z_{i-1};\boldsymbol{\theta}_n)$, is $\tilde{p}(\tilde{n}_{i-1}|z_{i-1};\boldsymbol{\theta}_n)$. To this aim, we define the vector

$$\mathbf{m} \triangleq \begin{pmatrix} z_i \\ \tilde{n}_i \end{pmatrix}.$$

Clearly, $\mathbf{m}$ is a complex normal random vector with pdf parameters

$$\boldsymbol{\mu}_{\mathbf{m}} \triangleq E\{\mathbf{m}\} = \begin{pmatrix} E\{z_i\} \\ E\{\tilde{n}_i\} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
$$\mathbf{C}_{\mathbf{m}} \triangleq \mathrm{cov}(\mathbf{m}) = E\left\{\begin{pmatrix} z_i z_i^* & z_i \tilde{n}_i^* \\ \tilde{n}_i z_i^* & \tilde{n}_i \tilde{n}_i^* \end{pmatrix}\right\} = \begin{pmatrix} \sigma_f^2 + \sigma_n^2 & \sigma_n^2 \\ \sigma_n^2 & \sigma_n^2 \end{pmatrix}$$

where in the last equality, we used the fact that $z_i = f_i + \tilde{n}_i$. The conditional pdf can now be derived using Bayes' rule, $p(A\,|\,B) = (p(A,B)/p(B))$ (in the following we omit the conditioning on $\boldsymbol{\theta}_n$, but all the pdfs assume given $\boldsymbol{\theta}_n$)

$$p(\tilde{n}_i|z_i) = \frac{p(\tilde{n}_i, z_i)}{p(z_i)} = \frac{\frac{1}{\pi^2|\mathbf{C}_{\mathbf{m}}|} e^{-\mathbf{m}^H \mathbf{C}_{\mathbf{m}}^{-1} \mathbf{m}}}{\frac{1}{\pi\left(\sigma_f^2+\sigma_n^2\right)} e^{-\frac{|z_i|^2}{\sigma_f^2+\sigma_n^2}}}$$
$$= \frac{1}{\pi \frac{\sigma_f^2 \sigma_n^2}{\sigma_f^2+\sigma_n^2}} \exp\left[-\frac{\sigma_f^2+\sigma_n^2}{\sigma_f^2\sigma_n^2}\left|\tilde{n}_i - \frac{\sigma_n^2}{\sigma_f^2+\sigma_n^2} z_i\right|^2\right]$$

where $|\cdot|$ denotes the determinant of a matrix, implying that $\tilde{n}_i|z_i \sim \mathcal{CN}(\sigma_n^2/(\sigma_f^2+\sigma_n^2)z_i, \sigma_f^2\sigma_n^2/(\sigma_f^2+\sigma_n^2))$. Thus, using (A.3), we can write the conditional expectation

$$E\{z_i|z_{i-1}\} = \alpha z_{i-1} - E\{\tilde{n}_{i-1}|z_{i-1}\} = \alpha z_{i-1} - \alpha \frac{\sigma_n^2 z_{i-1}}{\sigma_f^2+\sigma_n^2} = \frac{\alpha \sigma_f^2 z_{i-1}}{\sigma_f^2+\sigma_n^2}.$$

For evaluating the variance, we begin with

$$E\{|z_i|^2|z_{i-1}\} = E\{|\alpha z_{i-1} - \alpha \tilde{n}_{i-1} + \sqrt{1-\alpha^2} w_i + \tilde{n}_i|^2 | z_{i-1}\}.$$

Next, note that $\tilde{n}_i$ is independent of z_{i-1} and, therefore, also of $\tilde{n}_{i-1}|z_{i-1}$. We also note that w_i is independent of all the other terms, hence, all the cross terms, except those that involve both $\tilde{n}_{i-1}$ and z_{i-1} equal zero, leaving us with

$$\begin{aligned} E\{|z_i|^2|z_{i-1}\} &= \alpha^2|z_{i-1}|^2 - \alpha^2 z_{i-1}^* E\{\tilde{n}_{i-1}|z_{i-1}\} \\ &\quad - \alpha^2 z_{i-1} E\{\tilde{n}_{i-1}^*|z_{i-1}\} + \alpha^2 E\{|\tilde{n}_{i-1}|^2|z_{i-1}\} \\ &\quad + (1-\alpha^2)E\{|w_i|^2\} + E\{|\tilde{n}_i|^2\} \\ &= \alpha^2|z_{i-1}|^2 \left(\frac{\sigma_f^2}{\sigma_n^2+\sigma_f^2}\right)^2 + \alpha^2 \frac{\sigma_n^2\sigma_f^2}{\sigma_n^2+\sigma_f^2} \\ &\quad + (1-\alpha^2)\sigma_f^2 + \sigma_n^2. \end{aligned}$$

The conditional variance can now be readily written as

$$\mathrm{var}(z_i|z_{i-1}) = E\{|z_i|^2|z_{i-1}\} - |E\{z_i|z_{i-1}\}|^2 = \sigma_f^2 + \sigma_n^2 - \frac{\left(\alpha\sigma_f^2\right)^2}{\sigma_n^2+\sigma_f^2}. \quad \text{(A.5)}$$

Hence, the conditional pdf of $z_i|z_{i-1};\boldsymbol{\theta}_n$ is found to be

$$z_i|z_{i-1};\boldsymbol{\theta}_n \sim \mathcal{CN}\left(\frac{\alpha\sigma_f^2 z_{i-1}}{\sigma_n^2+\sigma_f^2}, \sigma_f^2+\sigma_n^2 - \frac{\left(\alpha\sigma_f^2\right)^2}{\sigma_n^2+\sigma_f^2}\right). \quad \text{(A.6)}$$

APPENDIX B
AN EXPLICIT EXPRESSION FOR THE ASYMPTOTIC VARIANCE OF THE PROPOSED ESTIMATOR

This appendix presents an expression for the asymptotic variance of the proposed estimator. The derivation of the asymptotic variance is complicated and beyond the scope and focus of this paper. It will appear in a following paper [22]. However, we chose to present the expression for the asymptotic variance here, in order to present a complete description of the estimator.

The asymptotic variance of the proposed estimator is given by

$$\mathrm{var}(\hat{\tau} - \tau_0) = \frac{1}{(N\tilde{J}(\tau_0))^2}(E_1 + E_2 + E_3 + E_4). \quad \text{(B.1)}$$

In the above expression

$$\tilde{J}(\tau_0) = \frac{2}{N} \frac{\tilde{\alpha}\left(\tilde{\sigma}_f^2\right)^2}{\left(\left(\tilde{\sigma}_f^2\right)^2 - \tilde{\alpha}^2\left(\tilde{\sigma}_f^2\right)^2\right)^2} r_f[1] \times \sum_{k=1}^{N-1} |s_k'^* s_k - s_{k-1}'^* s_{k-1}|^2 \quad \text{(B.2)}$$

where $r_f[k] = R_f^J[k]/\sigma_f^2, s_k \triangleq s_k(\tau)|_{\tau=\tau_0}, s_k' \triangleq (\partial s_k(\tau)/\partial\tau)|_{\tau=\tau_0}$, and $\tilde{\alpha}, \tilde{\sigma}_f^2$ are the mean of the estimates of the nuisance parameters produced by the estimator (which, due to their consistency, are asymptotically equal to their true values).

The expressions for E_1, E_2, E_3 and E_4 are

$$E_1 = 2W_1W_2 \sum_{l=1}^{N-1} |s_l' s_l^* - s_{l-1}' s_{l-1}^*|^2 \quad \text{(B.3)}$$

$$\begin{aligned} E_2 = W_1 \sum_{l=2}^{N-1} [&W_4 S(l,l-1)S(l-2,l-1) \\ &+ W_3 S(l,l-1)S(l-1,l-2) \\ &+ W_3 S(l-1,l)S(l-2,l-1) \\ &+ W_4 S(l-1,l)S(l-1,l-2)] \end{aligned}$$

$$\begin{aligned} E_3 = W_1 \sum_{l=1}^{N-2} [&W_4 S(l,l-1)S(l,l+1) \\ &+ W_3 S(l,l-1)S(l+1,l) \\ &+ W_3 S(l-1,l)S(l,l+1) \\ &+ W_4 S(l-1,l)S(l+1,l)] \end{aligned}$$

$$E_4 = W_1 \sum_{l=1}^{N-1} \sum_{m=1, m\neq l,l+1,l-1}^{N-1} [S(l,l-1) \times S(m-1,m)W_5(l,m) + S(l,l-1)S(m,m-1)W_6(l,m) + S(l-1,l)S(m-1,m)W_7(l,m) + S(l-1,l)S(m,m-1)W_5(l,m)]$$

where

$$W_1 = \left(\frac{\tilde{\alpha}\tilde{\sigma}_f^2}{\left(\tilde{\sigma}_f^2\right)^2 - \left(\tilde{\alpha}\tilde{\sigma}_f^2\right)^2} \right)^2 \quad \text{(B.4)}$$

$$W_2 = \left(\sigma_f^2 + \sigma_n^2\right)^2 - \left(\sigma_f^2\right)^2 r_f^2[1] \quad \text{(B.5)}$$

$$W_3 = \left(\sigma_f^2\right)^2 r_f^2[1] + \left(\sigma_f^2 + \sigma_n^2\right)\sigma_f^2 r_f[2] \quad \text{(B.6)}$$

$$W_4 = 2\left(\sigma_f^2\right)^2 r_f^2[1] \quad \text{(B.7)}$$

$$S(l,m) = s_l' s_l^* + s_m'^* s_m \quad \text{(B.8)}$$

$$W_5(l,m) = \left(\sigma_f^2\right)^2 r_f^2[1] + \left(\sigma_f^2\right)^2 r_f^2[l-m] \quad \text{(B.9)}$$

$$W_6(l,m) = \left(\sigma_f^2\right)^2 r_f^2[1] + \left(\sigma_f^2\right)^2 r_f[l-m-1] r_f[m-l-1] \quad \text{(B.10)}$$

$$W_7(l,m) = \left(\sigma_f^2\right)^2 r_f^2[1] + \left(\sigma_f^2\right)^2 r_f[l-m+1] r_f[m-l+1] \quad \text{(B.11)}$$

where σ_f^2 and σ_n^2 used in (B.5)–(B.11) are the true values of the nuisance parameters.

REFERENCES

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

[2] U. Mengali and A. D'Andrea, *Synchronization Techniques for Digital Receivers (Application of Communication Theory)*. New York: Plenum, 1997.

[3] T. Aulin and C.-E. Sundberg, "Synchronization properties of continuous-phase modulation," in *Proc. Global Telecommunications Conf.*, Miami, FL, Nov. 1982, pp. D7.1.1–D7.1.7.

[4] F. Patenaude, J. H. Lodge, and P. A. Galko, "Symbol timing tracking for continuous-phase modulation over fast flat-fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 615–626, Aug. 1991.

[5] K. Goethals and M. Moeneclaey, "Tracking performance of ML-oriented NDA symbol synchronizers for nonselective fading channels," *IEEE Trans. Commun.*, vol. 43, pp. 1179–1184, Feb.-Apr. 1995.

[6] H. L. Van Trees, *Detection, Estimation, and Modulation Theory—Part III*. New York: Wiley, 1971.

[7] S. S. Soliman and R. A. Scholtz, "Synchronization over fading dispersive channels," *IEEE Trans. Commun.*, vol. 36, pp. 499–505, Apr. 1988.

[8] F. Gini and G. B. Giannakis, "Frequency-offset estimation and symbol timing recovery in flat-fading channels: A cyclostationary approach," *IEEE Trans. Commun.*, vol. 46, pp. 400–411, Mar. 1998.

[9] M. Ghogho, A. Swami, and T. Durrani, "On blind carrier recovery in time-selective fading channels," in *Proc. 33rd Asilomar Conf. Signals, Systems, Computers*, vol. 1, Pacific Grove, CA, Oct. 1999, pp. 243–247.

[10] K. E. Scott and E. B. Olasz, "Simultaneous clock phase and frequency offset estimation," *IEEE Trans. Commun.*, vol. 43, pp. 2263–2270, July 1995.

[11] A. Swami, M. Ghogho, and T. Durrani, "Blind synchronization and Doppler spread estimation for MSK signals in time-selective fading channels," in *Proc. ICASSP*, vol. 5, Istanbul, Turkey, June 2000, pp. 2665–2668.

[12] J. M. Francos and B. Friedlander, "Bounds for estimation of multicomponent signals with random amplitude and deterministic phase," *IEEE Trans. Signal Processing*, vol. 43, pp. 1161–1172, May 1995.

[13] M. Ghogho, A. K. Nandi, and A. Swami, "Cramer–Rao bounds and maximum-likelihood estimation for random amplitude phase-modulated signals," *IEEE Trans. Signal Processing*, vol. 47, pp. 2905–2916, Nov. 1999.

[14] R. Dabora, J. Goldberg, and H. Messer, "Inherent limitations in data-aided time synchronization of continuous-phase modulation signals over time-selective fading channels," *IEEE Trans. Signal Processing*, vol. 50, pp. 1470–1482, June 2002.

[15] T. Aulin and C.-E. Sundberg, "Continuous phase modulation—Part I: Full response signaling," *IEEE Trans. Commun.*, vol. COM-29, pp. 196–209, Mar. 1981.

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

[17] F. Gini, M. Luise, and R. Reggiannini, "Cramer–Rao bounds in the parametric estimation of fading radiotransmission channels," *IEEE Trans. Commun.*, vol. 46, pp. 1390–1398, Oct. 1998.

[18] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation and Signal Processing*, ser. Telecommunications and Signal Processing. New York: Wiley, 1997.

[19] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[20] R. Dabora, "Time synchronization for continuous phase modulated signals over time selective fading channels," Master's thesis, Tel Aviv Univ., Tel Aviv, Israel, June 2000.

[21] P. Stoica, O. Besson, F. Vincent, and A. B. Gershman, "Approximate maximum-likelihood estimators for array processing in multiplicative noise environments," *IEEE Trans. Signal Processing*, vol. 48, pp. 2506–2518, Sept. 2000.

[22] R. Dabora, J. Goldberg, and H. Messer, *Performance of Mismatched Maximum Likelihood Estimation with Dependent Observations*, to be published.

Ron Dabora received the B.Sc. and M.Sc. degrees in electrical engineering in 1994 and 2000 respectively, from Tel-Aviv University, Tel-Aviv, Israel. He is currently working toward the Ph.D. degree at Cornell University, Ithaca, NY.

From 1994 to 2000, he was with the Signal Corps of Israel Defense Forces, and from 2000 to 2003, he was a member of the Algorithms Development Group, Millimetrix Broadband Networks, Israel.

Jason Goldberg (SM'99) received the B.S.E.E. and Ph.D. degrees in electrical engineering from Worcester Polytechnic Institute, Worcester, MA, and the University of London, London, U.K., in 1988 and 1993, respectively.

From 1993 to 1996, he was a Postdoctoral Researcher at Polytechnic University of Catalonia, Barcelona, Spain. From 1996 to 2003, he was a Staff Researcher at Tel-Aviv University, Tel-Aviv, Israel, in the area of statistical and array signal processing. He was also a member of the Research Group, Wireless Computing and Communications Division, Intel Corp., Petach Tikva, Israel, from 1999 to 2003, where his work focused on the design and analysis of baseband signal processing algorithms for wireless communications. He is currently a Member of the Technical Staff at Phonetact, Inc., Los Altos, CA.

Hagit Messer (F'00) is a Professor of Electrical Engineering at Tel Aviv University, Tel Aviv. Israel. She is an expert in statistical signal processing and applications to source localization, communication, and radar systems. She held international visiting positions at Philips, the University of Pennsylvania, Drexel University, Yale University, New Jersey Institute of Technology, and ENST-Paris. Parallel to her academic activities, she has been a consultant to several Israeli high-tech companies.

From February 2000 to March 2003, she was the Chief Scientist and the Head of the Directorate of Science at the Ministry of Science, Israel. In this capacity, she was in charge of the national strategic research program; the bilateral scientific relations with various countries (Germany, France, China, India, Korea and others); and the national science and society program. She was also responsible for the initiation and implementation of science-related operations within the government, including establishing and leading a national council for women and science; preparing a program for national priority to biotechnology; initiating a national council of bioethics; leading the preparation and the negotiation for the participation of Israel in the 6th European R&D framework program (FP6); dealing with the national broadband internet network; steering the participation of Israel in international science infrastructure such as CERN and SESAME. She is also devoted to promoting women in science and technology, and since its founding in 1999, she represents Israel in the EU Helsinki group for Women and Science.