Blind Channel and Carrier Frequency Offset Estimation Using Periodic Modulation Precoders

Erchin Serpedin, Antoine Chevreuil, Georgios B. Giannakis, *Fellow, IEEE,* and Philippe Loubaton, *Member, IEEE*

***Abstract*—Recent results have shown that blind channel estimators, which are resilient to the location of channel zeros, color of additive stationary noise, and channel order overestimation errors, can be developed for communication systems equipped with transmitter-induced cyclostationarity precoders. The present paper extends these blind estimation approaches to the more general problem of estimating the unknown intersymbol interference (ISI) and carrier frequency offset/Doppler effects using such precoders. An all-digital open-loop carrier frequency offset estimator is developed, and its asymptotic (large sample) performance is analyzed and compared to the Cramér–Rao bound (CRB). A subspace-based channel identification approach is proposed for estimating, in closed-form, the unknown channel, regardless of the channel spectral nulls. It is shown that compensating for the carrier frequency offset introduces no penalty in the asymptotic performance of the subspace channel estimator. Simulations are presented to corroborate the performance of the proposed algorithms.**

***Index Terms*—Communication channels, equalizers, estimation, frequency estimation, mobile communications.**

I. Introduction

RECENT works [7], [14], [28], [34] have shown that blind identification of single-input single-output (SISO) finite impulse response (FIR) communication channels can be achieved from second-order statistics of the received data without imposing any restriction on the channel zeros, color of additive noise, and channel order overestimation errors and without introducing excess redundancy in the transmitted sequence. The underlying channel identification approach relies on the output second-order cyclostationary (CS) statistics, which are induced by precoding the input symbol stream with a periodically time-varying precoder, which is referred to as a transmitter-induced cyclostationarity (TIC) precoder.

The purpose of the present paper is to extend these channel identifiability results to the more general problem of blind estimation of the carrier frequency offset and the channel, assuming knowledge only of the output CS-statistics induced at the transmitter by a TIC precoder that is implemented by modulating the input symbol stream with a strictly periodic and deterministic sequence. In this paper, algorithms for batch estimation and compensation for the unknown time-invariant channel and frequency offset effects from knowledge only of the received data are developed, and their asymptotic performance is analyzed and compared with the Cramér–Rao bound (CRB).

Blind equalization of frequency-selective channels with carrier frequency offset (FO) and/or Doppler effects is well motivated when it comes to compensating for local oscillator drifts. Thus far, only a limited number of approaches have been proposed for blind estimation of frequency–selective channels in the presence of carrier frequency offset. In [3], [19], and [22], the approach consists of two steps: First, the channel is equalized by minimizing a certain nonlinear cost function [constant modulus algorithm (CMA)], and then, the carrier phase is tracked from the equalized output. In general, these algorithms require a large number of samples to achieve convergence, and in the presence of residual inter-symbol interference (ISI) at the equalizer output, the performance of corresponding frequency offset estimators may degrade. In the present approach, the influence of residual ISI on the carrier offset estimate is alleviated, and consistency of the frequency offset estimators is guaranteed even in the presence of residual ISI. In addition, the present TIC framework enables development of a subspace channel identification algorithm that offers a closed-form expression for the unknown channel. Thus, the potential local minima or loss of convergence associated with [3], [19], and [22] in the presence of noise are also avoided. Furthermore, the TIC-based subspace approach is robust to channel overestimation errors, location of channel zeros, and color of additive stationary noise. It is shown that in the presence of FO, the proposed channel estimation approach achieves the same asymptotic performance as the corresponding subspace algorithm that assumes perfect carrier synchronization. In other words, compensating for FO does not introduce asymptotically any loss in the performance of the channel estimation approach. An additional feature of the proposed subspace channel estimation approach is guaranteed identifiability, irrespective of the channel zero locations. Simulations illustrate that equalization of channels with possibly spectral nulls within the TIC framework requires an order of magnitude less samples than a CMA-based equalizer. Recently, two additional approaches have been proposed for blind channel estimation and equalization of SISO FIR communication channels in the presence of carrier FO [], [32]. Based on iterative minimization of certain nonlinear functions, the algorithms proposed in [] and [32] may be computationally complex and prone to divergence in the presence of noise.

Manuscript received January 8, 1999; revised April 20, 2000. This work was supported by the Office of Naval Research under Grant N00014-93-1-0485. The associate editor coordinating the review of this paper and approving it for publication was Dr. Alle-Jan van der Veen.

E. Serpedin is with the Department of Electrical Engineering, Texas A&M University, College Station, TX 77843 USA.

A. Chevreuil and P. Loubaton are with the Laboratoire Systèmes de Communication, Université de Marne-la-Vallée, Marne-la-Vallée, France.

G. B. Giannakis is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: georgios@ece.umn.edu).

Publisher Item Identifier S 1053-587X(00)05969-9.

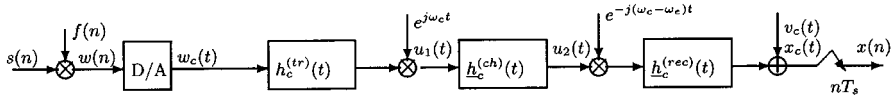

Fig. 1. Communication channel with frequency offset.

A different class of algorithms is proposed in [35] for blind separation/estimation of a nonconvolutive mixture of BPSK sources with residual carriers using the spatial diversity provided by an array of receiving antennas. The present paper assumes a different framework: a single-user channel with both convolutive and residual carrier effects and no spatial diversity condition but one that adopts a periodic precoder at the transmitter. Extension of the present single-user TIC approach to a multiuser scenario with both convolutive and residual carrier effects is possible when multiple users can be separated based on their distinct cyclic frequencies [9]. The proposed FO estimation setup can be viewed as an extension to frequency–selective channels of the FO estimation approach proposed for flat-fading channels in [18].

The rest of the paper is organized as follows. Section II introduces the communication channel setup and the concept of periodically time-varying TIC precoder. Sections III and IV propose channel and frequency offset estimation approaches based on the information provided by the received CS-statistics. In Sections V and VI, the asymptotic performance of the frequency offset and subspace channel estimators is investigated. Computer simulations are conducted in Section VII, and last, conclusions are drawn in Section VIII.

II. Preliminaries

We consider the general communication channel setup depicted in Fig. 1, where the independently and identically distributed (i.i.d.) input symbol stream $s(n)$ is modulated by the periodic sequence $f(n)$, $f(n) = f(n+P), \forall n$. The modulation of $s(n)$ with the periodic sequence $f(n)$ represents the nonredundant periodic TIC precoder, and its role is to induce cyclostationarity in the transmitted sequence $w(n) := f(n)s(n)$ (symbol $:=$ stands for "is defined as"). It can be easily checked that the second-order statistics of $w(n)$ are periodic. By defining the time-varying correlation of $w(n)$ at time n and lag τ as $c_{ww}(n;\tau) := \mathbb{E}w^*(n)w(n+\tau)$, (superscript $*$ stands for complex conjugation) , it follows that $c_{ww}(n;\tau) = c_{ww}(n+P;\tau) = |f(n)|^2\sigma_s^2\delta(\tau)$, where $\sigma_s^2 := \mathbb{E}|s(n)|^2$. Thus, $w(n)$ exhibits CS-statistics of period P. Sequence $w(n)$ is converted by the digital-to-analog (D/A) converter into the continuous-time waveform $w_c(t) := \sum_n w(n)\delta(t-nT_s)$, which is pulse shaped by the baseband transmit filter $h_c^{(tr)}(t)$ and modulated by the carrier $\exp(j\omega_c t)$. The resulting signal, which is denoted by $u_1(t)$, propagates through the unknown *bandpass* channel $\underline{h}_c^{(ch)}(t)$, (the underline _ denotes a bandpass representation), and at the receiver, it is demodulated with a carrier frequency offset ω_e, which is due to the local oscillator drifts. Defining the equivalent *baseband* components $h_c^{(ch)}(\tau) := \underline{h}_c^{(ch)}(\tau)e^{-j\omega_c\tau}$, $h_c^{(rec)}(\rho) := \underline{h}_c^{(rec)}(\rho)e^{-j\omega_e\rho}$, $h_c(t) := (h_c^{(rec)} \star h_c^{(ch)} \star h_c^{(tr)})(t)$, where $\star$ stands for convolution, the input/output (I/O) channel response can be expressed in the form

$$x_c(t) = e^{j\omega_e t}\sum_m w(m)h_c(t-mT_s) + v_c(t) \tag{1}$$

where the complex-valued noise $v_c(t)$ is assumed circularly Gaussian distributed with zero mean and independently distributed of $s(n)$. Sampling the continuous-time waveform $x_c(t)$ at the symbol rate $1/T_s$, the following discrete-time model is obtained from (1):

$$\begin{aligned} x(n) :=& x_c(nT_s) \\ =& e^{j\omega_e T_s n}\sum_m w(m)h_c((n-m)T_s) + v_c(nT_s). \end{aligned} \tag{2}$$

Considering $\theta_e := \omega_e T_s$, $h_c(t)$ of support LT_s, $h(n) := h_c(nT_s)$, and $v(n) := v_c(nT_s)$, from (2), we arrive at the equivalent discrete-time model shown in Fig. 2, which is described by the I/O relation

$$x(n) = e^{j\theta_e n}\sum_{l=0}^{L} h(l)w(n-l) + v(n). \tag{3}$$

In the next sections, a two-step approach is proposed for estimating the unknown channel and frequency offset from the information provided by the output CS-statistics. In the first step, the channel taps are recovered, within a scaling factor dependent on the unknown carrier frequency offset, using the second-order CS statistics of the output. In the second step, the carrier frequency offset is estimated from the second- or fourth-order CS statistics of the output and then used to remove the FO-dependent scaling factor present in the channel estimate.

III. Blind Channel Identification

The channel model with FO (3) can be rewritten in the equivalent form of a channel without FO, as is depicted in Fig. 3, but with an additional modulation superimposed on $s(n)$

$$x(n) = \sum_{l=0}^{L} g(l)u(n-l) + v(n) \tag{4}$$

where $g(l) := h(l)\exp(j\theta_e l)$, and $u(n) := w(n)\exp(j\theta_e n) = f(n)\exp(j\theta_e n)s(n)$. Since $c_{uu}(n;\tau) = c_{ww}(n;\tau)$, it follows that $u(n)$ exhibits CS statistics as well. Using (4), it can be easily shown that the output time-varying correlation $c_{xx}(n;\tau) := \mathbb{E}x^*(n)x(n+\tau) = c_{xx}(n+P;\tau)$ is also

periodic. Hence, it admits a Fourier series (FS) decomposition, whose coefficients are referred to as the cyclic correlations. The output cyclic correlation at cycle k and lag τ is given by [c.f. (4)]

$$\begin{aligned} &C_{xx}(k;\tau) \\ &\quad := \tfrac{1}{P}\sum_{n=0}^{P-1} c_{xx}(n;\tau)c^{-j2\pi kn/P} \\ &\quad = \sigma_s^2 F_2(k)\sum_{l} g^*(l)g(l+\tau)c^{-j2\pi kl/P} + \sigma_v^2\delta(\tau)\delta(k) \end{aligned} \tag{5}$$

where $F_2(k) := (1/P)\sum_{n=0}^{P-1}|f(n)|^2\exp(-j2\pi kn/P)$, $\sigma_s^2 := \mathbb{E}|s(n)|^2$, and $\sigma_v^2 := \mathbb{E}|v(n)|^2$. For a fixed cycle k, the cyclic spectrum corresponding to cycle k is defined as the Z-transform of the cyclic correlation sequence $\{C_{xx}(k;\tau)\}_\tau$, i.e., $S_{xx}(k;z) := \sum_\tau C_{xx}(k;\tau)z^{-\tau}$, and substituting, from (5), we arrive at

$$S_{xx}(k;z) = \sigma_s^2 F_2(k)G(z)G^*\left(z^{-1}e^{j2\pi k/P}\right) + \sigma_v^2\delta(k) \tag{6}$$

where $G(z) := \sum_{l=0}^{L} g(l)z^{-l} = H(z\exp(-j\theta_e))$, $G^*(z) := \sum_{l=0}^{L} g^*(l)z^{-l}$, and $H(z) := \sum_{l=0}^{L} h(l)z^{-l}$. Henceforth, only nonzero cycles $k \neq 0$ are considered, which implies that the contribution of the additive stationary noise $v(n)$ is cancelled out in (6). This key property is used next to derive a channel estimation approach that is resilient to the color of $v(n)$.

It is well-known that the transfer function $G(z)$ can be extracted as the greatest common divisor (gcd) of the family of cyclic spectra $S_{xx}(k;z)$, $k = 1,\cdots,P-1$, provided that the period of $f(n)$ satisfies $P > L+1$ [14]. By taking advantage of the structure of (6), a subspace channel estimation approach that is also robust to channel order overestimation errors may be developed, as in [8], [10], and [28]. In order to shorten the exposition of the subspace approach, we summarize here the main results presented in [8] and [10].

Associate with an arbitrary vector $\boldsymbol{a} := [a_1,\cdots,a_l]^T$ the $l(l-1)/2 \times l$ matrix $\mathcal{F}(\boldsymbol{a})$ defined as

$$\mathcal{F}(\boldsymbol{a}) := \left[\begin{array}{cccccc} -a_2 & a_1 & 0 & \cdots & 0 & 0 \\ -a_3 & 0 & a_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ -a_l & 0 & 0 & \cdots & 0 & a_1 \\ \hline 0 & -a_3 & a_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & -a_l & 0 & \cdots & 0 & a_2 \\ \hline \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \hline 0 & \cdots & \cdots & \cdots & -a_l & a_{l-1} \end{array}\right]. \tag{7}$$

Denote by $\mathbb{C}[z^{-1}]$ the set of polynomials in z^{-1} over the complex field $\mathbb{C}$ and by $\mathbb{C}^l[z^{-1}]$ the set of l-dimensional polynomial vectors $\boldsymbol{p}(z) := [p_1(z)\cdots p_l(z)]^T$, with $p_k(z) \in \mathbb{C}[z^{-1}]$, $1 \le k \le l$. Denoting by $d(z)$ the gcd of the set of polynomials $p_k(z)$, $1 \le k \le l$, the irreducible part of $\boldsymbol{p}(z)$ is defined as $\boldsymbol{p}_{ir}(z) := \boldsymbol{p}(z)/d(z) = [p_1(z)/d(z)\cdots p_l(z)/d(z)]^T$ [1]. By direct substitution, it follows that

$$\mathcal{F}(\boldsymbol{p}(z))\boldsymbol{p}_{ir}(z) = \boldsymbol{0} \quad \forall z. \tag{8}$$

The subspace approach that we borrow from [8] and [10] relies on (8). We collect all the cyclic spectra (6) corresponding to all nonzero cycles in the $(P-1)\times 1$ vector

$$\begin{aligned} \boldsymbol{s}_{xx}(z) := \big[&S_{xx}^*\left(1;z^{-1}\right), S_{xx}^*\left(2;z^{-1}\right),\cdots \\ &S_{xx}^*\left(P-1;z^{-1}\right)\big]^T. \end{aligned} \tag{9}$$

Due to (6), $\boldsymbol{s}_{xx}(z)$ in (9) can be factorized as

$$\boldsymbol{s}_{xx}(z) = \sigma_s^2\boldsymbol{g}_f(z)G^*\left(z^{-1}\right) \tag{10}$$

where $\boldsymbol{g}_f(z)$ is a $(P-1)\times 1$ vector given by

$$\begin{aligned} \boldsymbol{g}_f(z) := \Big[&F_2^*(1)G\left(ze^{-j2\pi/P}\right), F_2^*(2)G\left(ze^{-j4\pi/P}\right),\cdots \\ &F_2^*(P-1)G\left(ze^{-j2\pi(P-1)/P}\right)\Big]^T. \end{aligned} \tag{11}$$

Because the condition $P > L+1$ is assumed, $\boldsymbol{g}_f(z)$ is the irreducible part of $\boldsymbol{s}_{xx}(z)$. Taking into account (8), we thus have

$$\mathcal{F}\left(\boldsymbol{s}_{xx}(z)\right)\boldsymbol{g}_f(z) = \boldsymbol{0}. \tag{12}$$

Suppose that $G(z) := \sum_{l=0}^{\hat{L}} g(l)z^{-l}$, where $\hat{L}$ stands for an upper bound (or estimate) of the unknown channel order L. In order to write the polynomial vector equation (12) in the time domain, several additional notations are introduced next. The polynomial vector $\boldsymbol{s}_{xx}(z)$ given by (9) admits the power series expansion $\boldsymbol{s}_{xx}(z) = \sum_{\tau=-\hat{L}}^{\hat{L}} \boldsymbol{c}(\tau)z^{-\tau}$, where we set

$$\boldsymbol{c}(\tau) := [C_{xx}^*(1;-\tau), C_{xx}^*(2;-\tau),\cdots, C_{xx}^*(P-1;-\tau)]^T. \tag{13}$$

Let $\boldsymbol{e}(z)$ be the $(P-1)(P-2)/2\times 1$ vector defined as

$$\boldsymbol{e}(z) := \mathcal{F}\left(\boldsymbol{s}_{xx}(z)\right)\boldsymbol{g}_f(z) = \sum_{\tau=-\hat{L}}^{2\hat{L}} \boldsymbol{e}(\tau)z^{-\tau}. \tag{14}$$

The unknown polynomial vector $\boldsymbol{g}_f(z)$ can be written as $\boldsymbol{g}_f(z) = \sum_{\tau=0}^{\hat{L}} \boldsymbol{g}_{f,1}(\tau)z^{-\tau}$, where

$$\begin{aligned} &\boldsymbol{g}_{f,1}(\tau) \\ &\quad := \left[F_2^*(1)g(\tau)e^{j2\pi\tau/P}\cdots F_2^*(P-1)g(\tau)e^{j2\pi(P-1)\tau/P}\right]^T. \end{aligned}$$

In addition, define the vectors $\boldsymbol{c} := [\boldsymbol{c}^T(\hat{L}),\cdots,\boldsymbol{c}^T(-\hat{L})]^T$, $\boldsymbol{g}_{f,1} := [\boldsymbol{g}_{f,1}^T(0),\cdots,\boldsymbol{g}_{f,1}^T(\hat{L})]^T$, and $\tilde{\boldsymbol{g}} := [g(0)\cdots g(\hat{L})]^T$. Equation (14) can be rewritten in the equivalent matrix form

$$\boldsymbol{e} = \mathcal{T}(\mathcal{F}(\boldsymbol{c}))\boldsymbol{g}_{f,1} \tag{15}$$

where $\boldsymbol{e} := [\boldsymbol{e}(2\hat{L})\cdots\boldsymbol{e}(-\hat{L})]^T$, and $\mathcal{T}(\mathcal{F}(\boldsymbol{c}))$ denotes the $(3\hat{L}+1)(P-1)(P-2)/2\times(\hat{L}+1)(P-1)$ block Toeplitz matrix with the first column and row given by the vectors

$[\mathcal{F}(\boldsymbol{c}(\hat{L}))\cdots\mathcal{F}(\boldsymbol{c}(-\hat{L}))\mathbf{0}\cdots\mathbf{0}]^T$ and $[\mathcal{F}(\boldsymbol{c}(\hat{L}))\mathbf{0}\cdots\mathbf{0}]$, respectively. Parameterization of $\boldsymbol{g}_f(z)$ given by (11) provides, in the time domain, the relation

$$\boldsymbol{g}_{f,1} = \mathcal{P}\tilde{\boldsymbol{g}} \tag{16}$$

where the matrix $\mathcal{P}$ is block-diagonal with its qth diagonal entry given by

$$[\mathcal{P}]_{q,q} = \begin{bmatrix} F_2^*(1)e^{j2\pi(q-1)/P} \\ \vdots \\ F_2^*(P-1)e^{j2\pi(q-1)(P-1)/P} \end{bmatrix}. \tag{17}$$

Using (16), (15) can be written in the form

$$\mathcal{T}(\mathcal{F}(\boldsymbol{c}))\mathcal{P}\tilde{\boldsymbol{g}} = \mathbf{0}. \tag{18}$$

Identifiability of the subspace approach has been established in [7], [8], and [10] and is summarized by the following result.

Proposition 1 (Structured Subspace Identifiability Result): Assuming that the frequency offset θ_e is known, $P > L+1$, and $L \leq \hat{L} < L+P$, (18) admits a unique solution (within a scaling factor ambiguity) spanned by the $(\hat{L}+1)\times 1$ *vector* $[g(0), \cdots, g(L), 0\cdots 0]^T$.

Now, let us denote by $\boldsymbol{g} := [g(0)\cdots g(L)]^T$ and $\boldsymbol{h} := [h(0)\cdots h(L)]^T$ the impulse response vectors corresponding to the transfer functions $G(z)$ and $H(z)$, respectively. Since $G(z) = H(z\exp(-j\theta_e))$, it follows that

$$\boldsymbol{g} = \mathcal{D}\boldsymbol{h} \tag{19}$$

where $\mathcal{D}$ is a diagonal matrix defined as $\mathcal{D} := \text{diag}\{1\ \exp(j\theta_e)\cdots\exp(j\theta_e L)\}$. In practice, the cyclic correlations involved in $\boldsymbol{c}$ have to be estimated from a finite number of samples. A consistent estimate for the cyclic correlation at cycle k and lag τ is given by [12]

$$\hat{C}_{xx}(k;\tau) := \frac{1}{T}\sum_{n=0}^{T-1} x^*(n)x(n+\tau)e^{-j2\pi kn/P}. \tag{20}$$

Denote by $\hat{\boldsymbol{c}}$ an estimate for the vector of cyclic correlations $\boldsymbol{c}$, and assume that the channel order L is known. Plugging (19) into (18), it follows that

$$\boldsymbol{A}\boldsymbol{h} = \mathbf{0} \tag{21}$$

where $\boldsymbol{A} := \mathcal{T}(\mathcal{F}(\boldsymbol{c}))\mathcal{P}\mathcal{D}$. Because all blind channel identification approaches estimate the unknown channel within a scaling factor ambiguity, nonuniqueness of the channel estimate prevents us from computing the asymptotic performance. We remove the phase ambiguity by fixing $\boldsymbol{h}(1) = 1$. Thus, (21) can be rewritten as

$$\boldsymbol{A}_1\boldsymbol{h}_1 = \boldsymbol{b}_1 \tag{22}$$

where matrix $\boldsymbol{A}_1$ consists of the last L columns of $\boldsymbol{A}$, vector $-\boldsymbol{b}_1$ corresponds to the first column of $\boldsymbol{A}$, and $\boldsymbol{h}_1 := [h(1)h(2)\cdots h(L)]^T$. In addition, define the vectors of cyclic correlations $\boldsymbol{c}_{xx} := [C_{xx}(1;-L)\cdots C_{xx}(P-1;L)]^T$, $\hat{\boldsymbol{c}}_{xx} := [\hat{C}_{xx}(1;-L)\cdots\hat{C}_{xx}(P-1;L)]^T$ and the *normalized* asymptotic covariance matrix $\boldsymbol{\Sigma} := \lim_{T\to\infty} T(\hat{\boldsymbol{c}}_{xx} - \boldsymbol{c}_{xx})(\hat{\boldsymbol{c}}_{xx} - \boldsymbol{c}_{xx})^H$. By replacing the cyclic correlations with their sample estimates and the unknown FO θ_e with the estimate $\hat{\theta}_e$ [which is provided by (32) or (34)], (22) can be rewritten as

$$\hat{\boldsymbol{A}}_1\hat{\boldsymbol{h}}_1 = \hat{\boldsymbol{b}}_1. \tag{23}$$

Since transfer function $G(z)$ can be estimated without any knowledge of θ_e from the set of equations (18), the I/O relation (4) can be deconvolved, assuming the existence of perfect inverse $G^{-1}(z)$, to yield

$$x_e(n) := G^{-1}(z)x(n) = e^{j\theta_e n}f(n)s(n) + G^{-1}(z)v(n). \tag{24}$$

We observe from (19) and (24) that recovery of $h(n)$ and $s(n)$ requires estimation of θ_e. Before proceeding any further, we suppose that estimation and deconvolution of $G(z)$ have been performed, and we concentrate next on the frequency offset estimation problem relying on $x_e(n)$.

IV. Frequency Offset Estimation

Estimating θ_e from $x_e(n)$ in (24) amounts to retrieving a complex exponential in multiplicative noise [$s(n)$ is random] and additive noise $G^{-1}(z)v(n)$: a problem originally addressed in [16] and [36]. Two different frequency offset estimators proposed herein rely on the results of [16] and [36]. The first FO-estimator applies to input symbol constellations that satisfy the moment condition $\mathbb{E}s^2(n) \neq 0$. For such constellations, recovery of θ_e is shown to be possible from the output second order CS statistics. Real-valued constellations (e.g., BPSK, PAM) are typical examples of input constellations for which the moment condition $\mathbb{E}s^2(n) \neq 0$ holds. The second estimator applies to input constellations that satisfy the moment conditions $\mathbb{E}s^2(n) = \mathbb{E}s^3(n) = 0$, $\mathbb{E}s^4(n) \neq 0$. For such constellations, it is shown that the output fourth order CS statistics provide enough information for estimating θ_e. Complex–valued constellations [e.g., QAM(4) $\div$ QAM(32)] are examples of input constellations for which these moment conditions are satisfied.

A. BPSK Constellations

It is known that for real-valued input constellations $s(n)$ (e.g., BPSK, PAM), which satisfy $\mathbb{E}s^2(n) \neq 0$, the second-order *conjugate* cyclic correlations of $x_e(n)$ allow recovery of θ_e [16], [36]. Consider that $s(n)$ is a BPSK sequence, and define the conjugate time-varying correlation

$$\begin{aligned}\tilde{c}_{x_e x_e}(n;\tau) &:= \mathbb{E}x_e(n)x_e(n+\tau) \\ &= \sigma_s^2 e^{j\theta_e(2n+\tau)}f^2(n)\delta(\tau) + \tilde{c}_{vv}(\tau)\end{aligned} \tag{25}$$

where $\tilde{c}_{vv}(\tau) := \mathbb{E}[G^{-1}(z)v(n)][G^{-1}(z)v(n+\tau)]$. Being almost periodically varying, the generalized Fourier series coefficient of $\tilde{c}_{x_e x_e}(n;0)$, which is termed the conjugate cyclic correlation, is given by [c.f. (25)]

$$\begin{aligned}\tilde{C}_{x_e x_e}(\alpha;0) &:= \lim_{T\to\infty}\frac{1}{T}\sum_{n=0}^{T-1}\tilde{c}_{x_e x_e}(n;0)e^{-j\alpha n} \\ &= \sigma_s^2\tilde{F}_2(\alpha - 2\theta_e) + \tilde{c}_{vv}(0)\delta(\alpha)\end{aligned} \tag{26}$$

where $\tilde{F}_2(\alpha) := \lim_{T\to\infty}(1/T)\sum_{n=0}^{T-1} f^2(n)\exp(-j\alpha n)$. Since $f(n)$ is strictly periodic, $\tilde{F}_2(\alpha)$ consists of Kronecker deltas located at the harmonics $2\pi k/P$, with k being an integer;

i.e., $\tilde{F}_2(\alpha) = \sum_k \tilde{F}_2(k)\delta(\alpha - 2\pi k/P)$. From (26), it follows that $\tilde{C}_{x_e x_e}(\alpha; 0)$ also consists of Kronecker deltas (or spectral lines) located at cycles $2\theta_e + 2\pi k/P$, where k is an integer. An estimate of the FO θ_e can be obtained by measuring the location of these spectral lines. Consider the following nonlinear least-squares (NLS) estimator

$$\hat{\boldsymbol{\theta}} := \arg\min_{\boldsymbol{\theta}} J(\boldsymbol{\theta}) \tag{27}$$

$$J(\boldsymbol{\theta}) := \frac{1}{T}\sum_{n=0}^{T-1}\left| x_e^2(n) - \sum_{k=0}^{P-1} \alpha_k e^{j\phi_k} e^{j(2\theta_e + (2\pi k/P))n}\right|^2 \tag{28}$$

where $\boldsymbol{\theta} := [\alpha_0 \phi_0 \cdots \alpha_{P-1}\phi_{P-1} 2\theta_e]^T$, with the real-valued variables $\alpha_k > 0$, $\phi_k \in [0, 2\pi)$, $k = 0, \cdots, P-1$, and $|2\theta_e| < \pi/P$, where the last inequality is imposed to ensure identifiability of FO θ_e.

Due to the nonlinear minimization involved by the NLS-estimator (27) and (28), it is of interest to obtain a computationally efficient algorithm. Straightforward manipulations show that[1]

$$\begin{aligned}\frac{\partial J(\boldsymbol{\theta})}{\partial \alpha_k} = 0 \Rightarrow \alpha_k &= \frac{1}{T}\,\mathrm{Re}\left\{\sum_{n=0}^{T-1} x_e^2(n) e^{-j\phi_k} e^{-j(2\theta_e + (2\pi k/P))n}\right\} + o(1) \\ \frac{\partial J(\boldsymbol{\theta})}{\partial \phi_k} = 0 \Rightarrow \frac{1}{T}&\left\{\sum_{n=0}^{T-1} x_e^2(n) e^{-j\phi_k} e^{-j(2\theta_e + (2\pi k/P))n}\right\} \\ &= \frac{1}{T}\left\{\sum_{n=0}^{T-1} x_e^{2*}(n) e^{j\phi_k} e^{j(2\theta_e + (2\pi k/P))n}\right\} + o(1).\end{aligned} \tag{29, 30}$$

Substituting (29) into (28) and taking into account (30), it follows that

$$J(\boldsymbol{\theta}) = \frac{1}{T}\sum_{n=0}^{T-1} |x_e(n)|^4 - \sum_{k=0}^{P-1}\left|\frac{1}{T}\sum_{n=0}^{T-1} x_e^2(n) e^{-j(2\theta_e + (2\pi k/P))n}\right|^2 + o(1). \tag{31}$$

Thus, asymptotically, the NLS-estimator coincides with the estimator (see also [36])

$$\hat{\theta}_e := \arg\min_{\theta_e} \sum_{k=0}^{P-1}\left|\frac{1}{T}\sum_{n=0}^{T-1} x_e^2(n) e^{-j(2\theta_e + (2\pi k/P))n}\right|^2. \tag{32}$$

The FO-estimator (32) can be implemented efficiently by considering the fast Fourier transform (FFT) of the squared data. Zero-padding of the data to a sufficiently large number of points (T_{zp}) prior to the FFT is necessary so that the frequency bins are small enough to allow accurate estimation of the FO. For the FFT bin to be compatible with this limit, it is necessary to ensure a zero-padding on the order of $1/T_{zp} < \sqrt{\mathrm{CRB}(\theta_e)}$ (see e.g., [31, pp. 147–150]).

[1]By $a(T) = O(T^k)$, we mean that $\lim_{T\to\infty} a(T)/T^K = c$, where the constant c satisfies $0 < |c| < \infty$. The notation $a(T) = o(T^k)$ stands for $\lim_{T\to\infty} a(T)/T^k = 0$.

B. QAM Constellations

A similar carrier frequency-offset estimator can be developed when the input belongs to a QAM constellation. Suppose now that $s(n)$ is drawn from a QAM constellation that satisfies the moment conditions $\mathbb{E}s^2(n) = 0$, $\mathbb{E}s^3(n) = 0$, and $\mathbb{E}s^4(n) \neq 0$. Define the fourth-order time-varying correlation $\tilde{c}_{4,x_e x_e}(n;\tau) := \mathbb{E}x_e^2(n)x_e^2(n+\tau)$. It can be easily checked that $\tilde{c}_{4,x_e x_e}(n;\tau) = \mathbb{E}s^4(n)\exp(j(4\theta_e n + 2\theta_e\tau))f^4(n)\delta(\tau) + \tilde{c}_{4,vv}(n;\tau)$, with $\tilde{c}_{4,vv}(n;\tau) := \mathbb{E}[G^{-1}(z)v^2(n)][G^{-1}(z)v^2(n+\tau)]$. The generalized Fourier series coefficient of $\tilde{c}_{4,x_e x_e}(n;0)$ can be expressed as

$$\begin{aligned}\tilde{C}_{4,x_e x_e}(\alpha;0) &:= \lim_{T\to\infty}\frac{1}{T}\sum_{n=0}^{T-1}\tilde{c}_{4,x_e x_e}(n;0)e^{-j\alpha n} \\ &= \mathbb{E}s^4(n)\tilde{F}_4(\alpha - 4\theta_e) + \tilde{c}_{4,vv}(n;0)\delta(\alpha)\end{aligned} \tag{33}$$

with $\tilde{F}_4(\alpha) := \lim_{T\to\infty}(1/T)\sum_{n=0}^{T-1} f^4(n)\exp(-j\alpha n)$. Since $\tilde{F}_4(\alpha)$ consists of a sequence of Kronecker deltas located at harmonic cycles $2\pi k/P$, $\tilde{C}_{4,x_e x_e}(\alpha;0)$ also consists of a sequence of Kronecker deltas located at $4\theta_e + 2\pi k/P$, $k = 0, \cdots, P-1$. Similar to (32), an FFT-based estimator can be obtained as

$$\hat{\theta}_e := \arg\min_{\theta_e}\sum_{k=0}^{P-1}\left|\frac{1}{T}\sum_{n=0}^{T-1} x_e^4(n)e^{-j(4\theta_e + (2\pi k/P))n}\right|^2. \tag{34}$$

As before, the assumption $|4\theta_e| < \pi/P$ is necessary to ensure uniqueness of the FO $\hat{\theta}_e$ in (34). In the next section, the asymptotic performance of FO-estimator (32) is established. Because analysis of the FO-estimator (34) entails similar steps, we omit its derivation. We conclude this section by summarizing the main steps of the proposed channel and FO estimation algorithm.

Step 1) Compute the cyclic correlation estimates $\hat{C}_{xx}(k;\tau)$ [using e.g., (20)].
Step 2) Find $\hat{\boldsymbol{g}}$ from the sample estimate of (18).
Step 3) Equalize the channel (4).
Step 4) Find $\hat{\theta}_e$ using (32) or (34).
Step 5) Find $\hat{\boldsymbol{h}}$ using (19).

V. Asymptotic Performance of FO-Estimators

We will assume that the input $s(n)$ is BPSK (± 1) and that the zero-mean additive noise $v(n)$ satisfies the mixing condition

$$\sum_{n_1}\cdots\sum_{n_K}|c_{v\cdots v}(n_1, \cdots, n_K)| < \infty, \qquad K = 1, 2, \cdots \tag{35}$$

where $c_{v\cdots v}(n_1, \cdots, n_K) := \mathrm{cum}\{v(n+n_1), \cdots, v(n+n_K), v(n)\}$ stands for the $(K+1)$st-order cumulant of $v(n)$. Condition (35) refers to the absolute summability of cumulants

of any order and is a reasonable assumption in practice since it is satisfied by all time series of weak memory, i.e., an asymptotically vanishing span of dependence of the available time series samples [4, pp. 8, 25–27]. Assumption (35) will prove useful in establishing the consistency of our estimation algorithms and in facilitating calculation of the asymptotic performance of the proposed estimators. From (24), we obtain for $n = 0, \cdots, T-1$ [since $s(n)$ is assumed BPSK(± 1), $s^2(n) = 1$]

$$x_e^2(n) = e^{j2\theta_e n} f^2(n) s^2(n) + \epsilon(n) \tag{36}$$

$$\epsilon(n) := 2e^{j\theta_e n} f(n) s(n) \left[G^{-1}(z) v(n)\right] + \left[G^{-1}(z) v(n)\right]^2. \tag{37}$$

Letting $F_2(k) := \alpha_k \exp(j\phi_k)$ and substituting $f^2(n) = \sum_{k=0}^{P-1} F_2(k) \exp(j2\pi kn/P)$ into (36), we obtain

$$x_e^2(n) = \sum_{k=0}^{P-1} \alpha_k e^{j(((2\pi k/P)+2\theta_e)n+\phi_k)} + \epsilon(n).$$

Since estimators (32) and (27) coincide asymptotically, it suffices to establish the asymptotic performance of the NLS-estimator (27). We follow the lines of proof presented in [5], [20], and [21]. After some lengthy calculations, whose details are presented in Appendix A, the following expression is obtained for the FO-estimator's asymptotic variance, as shown in (38) at the bottom of the page, with $\alpha_k \exp(j\phi_k) := F_2(k)$, $\alpha_k > 0$, and $S_{\epsilon\epsilon}(k, \exp(j\omega))$ standing for the cyclic spectrum of CS-noise $\epsilon(n)$ evaluated at cycle k and frequency ω.

Relation (38) does not allow any simple interpretation for the asymptotic variance, except the fact that it depends inversely proportional on the sum of squares of the magnitude of all spectral lines and is directly proportional to the cyclic spectra of additive CS noise $\epsilon(n)$. However, it is important to recognize the rapid rate of convergence of the FO estimator ($O(T^{-3})$). This observation will prove useful in the next section, where it is shown that the operation of compensating for the FO does not introduce any penalty in the asymptotic performance of the subspace channel estimator. This fast rate of convergence ($O(T^{-3})$) of the proposed FO estimator will be maintained even in the presence of residual ISI effects; the only modification in (38) consists of including in the cyclic spectra $S_{\epsilon\epsilon}(k, \exp(j\omega))$ of the contributions due to residual ISI left uncompensated.

To benchmark performance of our FO estimator, the performance of the stochastic maximum likelihood method is assessed here by deriving its CRB. The vector of unknown parameters is denoted by $\chi := [\sigma_v^2 \quad \theta_e]^T$, and the unknown but random transmitted symbols are grouped together in the vector $\boldsymbol{s} := [s(0) \cdots s(T-1)]^T$. Computation of the CRB requires knowledge of the second-order statistics of additive noise $v(n)$ (which may not be available). To simplify the calculations, additive noise $v(n)$ is assumed white and normally distributed. The CRB for an unbiased estimator $\hat{\chi}$ is bounded below by the inverse of the Fisher information matrix (FIM) $\boldsymbol{J}(\chi)$:

$$\mathbb{E}\left\{(\hat{\chi} - \chi)(\hat{\chi} - \chi)^T\right\} \geq \boldsymbol{J}^{-1}(\chi)$$

where $\boldsymbol{J}(\chi) = -\mathbb{E}\{\partial^2 \log f_\chi(\boldsymbol{x})/\partial\chi_k \partial\chi_l\}$, $k, l = 1, 2$, $\boldsymbol{x} := [x(0)x(1)\cdots x(T-1)]^T$, and χ_k denotes the kth entry of χ, and $f_\chi(\boldsymbol{x})$ stands for the likelihood function of $\boldsymbol{x}$. Since input $s(n)$ is assumed unknown but random, the conditional likelihood $f_\chi(\boldsymbol{x}|\boldsymbol{s})$ has to be averaged over all input sequences $\boldsymbol{s}$: $f_\chi(\boldsymbol{x}) = \mathbb{E}_{\boldsymbol{s}}[f_\chi(\boldsymbol{x}|\boldsymbol{s})]$. The exact CRB is given by

$$\boldsymbol{J}(\chi) = -\mathbb{E}\left\{\frac{\partial^2 \log \mathbb{E}_{\boldsymbol{s}}\left[f_\chi(\boldsymbol{x}|\boldsymbol{s})\right]}{\partial\chi^2}\right\}. \tag{39}$$

Since the evaluation of $\mathbb{E}_{\boldsymbol{s}}[f_\chi(\boldsymbol{x}|\boldsymbol{s})]$ is computationally intractable, it is common to adopt (see, e.g., [], [13], and [17]) a looser bound, which is referred to as the stochastic CRB, which is not as tight as the exact CRB, but it is computationally easier to evaluate. Due to the concavity of the log function and Jensen's inequality, we obtain from (39) the following valid CRB bound:

$$\boldsymbol{J}(\chi) \leq -\mathbb{E}_{\boldsymbol{s}}\mathbb{E}\left\{\frac{\partial^2 \log\left[f_\chi(\boldsymbol{x}|\boldsymbol{s})\right]}{\partial\chi^2}\right\}. \tag{40}$$

Direct calculations show that $\mathbb{E}\{\partial^2 \log f_\chi(\boldsymbol{x}|\boldsymbol{s})/\partial\sigma_v^2\partial\theta_e\} = 0$, and $-\mathbb{E}_{\boldsymbol{s}}\mathbb{E}\{\partial^2 \log f_\chi(\boldsymbol{x}|\boldsymbol{s})/\partial^2\theta_e\} = \sum_{n=L}^{T-1}\sum_{l=0}^{L}|h(l)|^2(n-l)^2|f(n-l)|^2/\sigma_v^2$. Thus, we obtain

$$\mathbb{E}\left\{\left(\hat{\theta}_e - \theta_e\right)^2\right\} \geq \frac{\sigma_v^2}{\sum_{n=L}^{T-1}\sum_{l=0}^{L}|h(l)|^2(n-l)^2|f(n-l)|^2} \approx \frac{\sigma_v^2}{\left(\sum_{l=0}^{L}|h(l)|^2\right)\left(\sum_{n=1}^{T-1} n^2|f(n)|^2\right)}. \tag{41}$$

VI. Asymptotic Performance of the Subspace Channel Estimator

The asymptotic performance of the least-squares (LS) channel estimate $\hat{\boldsymbol{h}}_1$ in (23) can be established using standard approaches (see, e.g., [24, pp. 96 and 97]) and is given by the following result.

$$\lim_{T\to\infty} T^3\left(\hat{\theta}_e - \theta_e\right)^2 = \frac{3\sum_{k=0}^{P-1}\sum_{l=0}^{P-1}\alpha_k\alpha_l \operatorname{Re}\left\{e^{j(\phi_k-\phi_k)}S_{\varepsilon\varepsilon}\left(k-l;\, e^{j((2\pi k/P)+2\theta_e)}\right)\right\}}{2\left(\sum_{k=0}^{P-1}\alpha_k^2\right)^2}. \tag{38}$$

Theorem 1: The asymptotic estimation error $\sqrt{T}(\hat{\boldsymbol{h}}_1 - \boldsymbol{h}_1)$ has a limiting zero-mean complex Gaussian distribution with asymptotic covariance

$$\lim_{T\to\infty} T\mathbb{E}\left[\left(\hat{\boldsymbol{h}}_1 - \boldsymbol{h}_1\right)\left(\hat{\boldsymbol{h}}_1 - \boldsymbol{h}_1\right)^H\right] = \left(\boldsymbol{A}_1^H\boldsymbol{A}_1\right)^{-1}\boldsymbol{A}_1^H\boldsymbol{D}\boldsymbol{\Sigma}\boldsymbol{D}^H\boldsymbol{A}_1\left(\boldsymbol{A}_1^H\boldsymbol{A}_1\right)^{-1} \quad (42)$$

where $\boldsymbol{D}$ is the matrix with its lth column given by

$$\boldsymbol{D}(:, l) = \frac{\partial \boldsymbol{b}_1}{\partial \boldsymbol{c}_l} - \frac{\partial \boldsymbol{A}_1}{\partial \boldsymbol{c}_l}\boldsymbol{h}_1 \quad (43)$$

and c_l denotes the lth entry of cyclic vector $\boldsymbol{c}_{xx}$.

Proof: See Appendix D.

Next, we show that the same asymptotic covariance matrix is obtained for the subspace channel estimator in the presence and/or absence of carrier frequency offset whenever the input source is circularly distributed. We first show that the asymptotic mean-square error of the channel estimate $\hat{\boldsymbol{h}}_1$, corresponding to the FO-channel model of Fig. 2, is equal to the asymptotic mean-square error of the subspace estimate corresponding to the channel model depicted in Fig. 4, which assumes perfect carrier frequency offset synchronization but the channel $g(n)$ instead of $h(n)$. Keeping in mind (42) and that the FO-channel model of Fig. 2 can be represented as in Fig. 3, it suffices to show that the asymptotic covariance matrix $\boldsymbol{\Sigma}$ is the same for the channels represented in Figs. 3 and 4. From the proof of Proposition 3, it will follow that the asymptotic covariance $\boldsymbol{\Sigma}$ of the cyclic correlation coefficients is not dependent on the modulating sequence $\exp(j\theta_e n)$ that is present in $u(n)$ [$u(n) = f(n)\exp(j\theta_e n)s(n)$; see Fig. 3]. This claim relies on the circularity assumption on $s(n)$.

Proposition 3: Assuming that $s(n)$ is circularly distributed, the asymptotic covariance matrix $\boldsymbol{\Sigma}$ does not depend on the modulating sequence $\exp(j\theta_e n)$ *present in* $u(n)$. *For given cycles* $k_1, k_2, 1 \le k_1, k_2 \le P-1$, the following asymptotic result holds:

$$\lim_{T\to\infty} T\,cov\left(\hat{C}_{xx}(k_1;\tau_1), \hat{C}_{xx}(k_2;\tau_2)\right) = \xi_1 + \xi_2 \quad (44)$$

where

$$\xi_1 = \sum_{\substack{\ell_1,\ell_2=0 \\ \ell_1-\ell_2\equiv k_1-k_2}}^{P-1} e^{j2\pi(\ell_1/P)\tau_2}\int_0^1 S_{xx}\left(\ell_1; e^{j2\pi f}\right) \cdot S_{xx}^*\left(\ell_2; e^{j2\pi(f-(k_1/P))}\right)e^{j2\pi f(\tau_1-\tau_2)}\,df \quad (45)$$

and

$$\xi_2 = \theta(k_1,k_2)\int_0^1\int_0^1 G\left(e^{j2\pi\nu_1}\right)G^*\left(e^{j2\pi(\nu_1-(k_1/P))}\right) \cdot G^*\left(e^{j2\pi\nu_2}\right)G\left(e^{j2\pi(\nu_2-(k_2/P))}\right) \cdot e^{j2\pi(\nu_1\tau_1-\nu_2\tau_2)}\,d\nu_1\,d\nu_2 \quad (46)$$

where $\theta(k_1, k_2) := \gamma_{4s}F_4(k_1 - k_2)$, γ_{4s} stands for the kurtosis of $s(n)$, and $F_4(k) := (1/P)\sum_{n=0}^{P-1}|f(n)|^4\exp(-j2\pi kn/P)$.

Proof: See Appendix E.

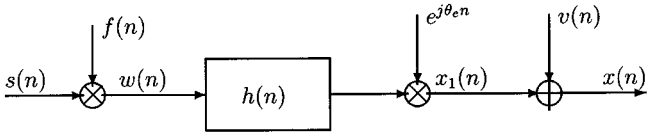

Fig. 2. Equivalent discrete-time model.

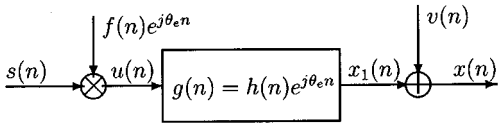

Fig. 3. Equivalent discrete-time model: Without FO.

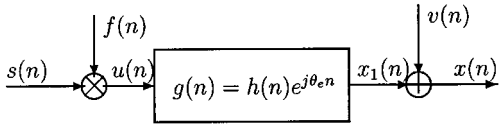

Fig. 4. Equivalent asymptotic model.

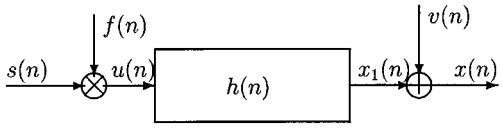

Fig. 5. Equivalent asymptotic model.

We show next that the same asymptotic distribution holds for the subspace channel estimates corresponding to the channels depicted in Figs. 4 and 5. In Fig. 4, the unknown FO θ_e is supposed to be given by the FO estimator (34). Let $S_{xx}^1(k; z)$ and $S_{xx}^0(k; z)$ denote the output cyclic spectra corresponding to the channels depicted in Figs. 4 and 5, respectively. Since $G(z) = H(z\exp(-j\theta_e))$, it follows that $S_{xx}^1(k; z) = S_{xx}^0(k; z\exp(-j\theta_e))$. Thus, an estimate of $S_{xx}^1(k; z)$ can be obtained from an estimate of $S_{xx}^0(k; z\exp(-j\theta_e))$ using the relation

$$\hat{S}_{xx}^1(k; z) := \hat{S}_{xx}^0\left(k; ze^{-j\hat{\theta}_e}\right), \quad \forall k = 1, \cdots, P-1 \quad (47)$$

where $\hat{S}_{xx}^0(k; z)$ stands for a consistent estimate of $S_{xx}^0(k; z)$, and $\hat{\theta}_e$ denotes the estimate (34) of θ_e. Let $\hat{\boldsymbol{h}}^1$ and $\hat{\boldsymbol{h}}^0$ denote the subspace channel estimates corresponding to the channels represented in Figs. 4 and 5, respectively, which are obtained using (21). It then holds that

$$\hat{\boldsymbol{A}}^1\hat{\boldsymbol{h}}^1 = \boldsymbol{0} \quad (48)$$

$$\hat{\boldsymbol{A}}^0\hat{\boldsymbol{h}}^0 = \boldsymbol{0}. \quad (49)$$

Interestingly, (47) implies the following link between $\hat{\boldsymbol{A}}^1$ and $\hat{\boldsymbol{A}}^0$:

$$\hat{\boldsymbol{A}}^1 = \hat{\boldsymbol{E}}\hat{\boldsymbol{A}}^0 \quad (50)$$

where

- $\hat{\boldsymbol{E}}$ diagonal matrix: $\hat{\boldsymbol{E}} := \boldsymbol{I} \otimes \boldsymbol{E}_0$;
- $\boldsymbol{I}$ $(P-1)(P-2)/2 \times (P-1)(P-2)/2$ identity matrix, $\boldsymbol{E}_0 := \text{diag}\{\exp(-jL\hat{\theta}_e)\exp(-j(L-1)\hat{\theta}_e)\cdots\exp(j2L\hat{\theta}_e)\}$;
- $\otimes$ Kronecker product.

Plugging (50) into (48), we obtain

$$\hat{\boldsymbol{E}}\hat{\boldsymbol{A}}^0\hat{\boldsymbol{h}}^1 = 0. \tag{51}$$

Comparing (49) and (51) and repeating the arguments used in the proof of Theorem 1, it follows that the channel estimates $\hat{\boldsymbol{h}}^1$ and $\hat{\boldsymbol{h}}^0$ have the same asymptotic distribution. We have thus established the following theorem.

Theorem 2: Assuming the input is circularly distributed, compensating for the carrier FO does not introduce any penalty in the asymptotic performance of the subspace channel estimator.

An intuitive explanation of this result may be given on the grounds that the frequency estimator has a much faster rate of convergence $(O(T^{-3}))$ than the subspace channel estimator $(O(T^{-1}))$. Thus, for sufficiently large $T \gg 1$, the FO estimate may be considered to be equal to its true value and can be considered compensated exactly without affecting the channel estimate.

It is interesting to note that in general, two frequency offset estimation setups are possible: post-equalization FO estimator, which estimates the FO θ_e from knowledge of equalized samples $x_e(n)$ (the setup assumed in this paper) and pre-equalization FO estimator, which estimates the FO θ_e directly from the received samples $x(n)$. Pre-equalization frequency offset estimators can be similarly derived by replacing $x_e(n)$ with $x(n)$ in the previous developments, e.g., the estimator (32) takes the form

$$\hat{\theta}_e := \arg\min_{\theta_e} \sum_{k=0}^{P-1} \left| \tfrac{1}{T} \sum_{n=0}^{T-1} x^2(n) e^{-j(2\theta_e + (2\pi k/P))n} \right|^2. \tag{52}$$

A natural question that arises is: Which one of the two frequency offset/channel estimation setups is preferable in practice? We discuss separately the two estimation problems: frequency offset and channel. From a frequency offset estimation viewpoint, it is clear that asymptotically ($T \to \infty$) and for large signal-to-noise ratios (SNR's) $\sigma_v^2 \to 0$, the post-equalization setup is preferable since the asymptotic variance (38) attains the value zero. This is not the case for the pre-equalization setup, where even in the absence of additive noise $v(n)$, the asymptotic variance (38) is nonzero due to residual channel ISI effects that are left uncompensated. At very low SNR's and in the presence of channel spectral nulls, it is expected that pre-equalization will cause noise amplification. This may affect the variance of $\hat{\theta}_e$ and the bit error rate. In such a scenario, the pre-equalization estimation of FO θ_e may appear as an alternative. Due to the very intricate expressions that are involved, this analysis is reported in the companion paper [11], where it is shown that estimation and compensation of the carrier frequency offset can be performed first, assuming the presence of an unknown frequency selective channel and with no loss in performance relative to methods where the channel is first pre-equalized, and then, the frequency offset is compensated afterwards.

From a channel estimation viewpoint, it is important to note that both pre- and post-equalization FO estimators are asymptotically ($T \to \infty$) similar. This is because the pre-equalization FO estimators also converge also at a rate $O(T^{-3})$ and by repeating the arguments involved in proving Theorem 2, one reaches the conclusion that asymptotically pre-equalization channel estimation schemes behave similarly, as if there is no carrier FO. However, the comparison is difficult to perform when only a finite number of samples is available. This problem is beyond the scope and length of the present paper and is to be reported in a future paper.

In the next section, several simulation experiments are conducted in order to test the performance of the proposed algorithms.

VII. Simulations

Throughout our simulations, we considered a channel with the baseband channel impulse response $\boldsymbol{h} := [0.53 + 0.07j, -0.24 - 0.23j, -0.54 - 0.32j, 0.11 + 0.44j, -0.036 - 0.099j]^T$, frequency offset $\theta_e = \pi/30$, and periodic precoder $\{f(n)\}_{n=1}^{P} := [\sqrt{\rho}, \cdots, \sqrt{\rho}, \sqrt{P(1-\rho)+\rho}]$, $\rho = 0.5878$, $P = 5$. The input $s(n)$ was drawn from an equiprobable QPSK or BPSK constellation, and the additive noise $v(n)$ was white and normally distributed, with zero mean. As channel estimation performance measures, the normalized root-mean-square error (RMSE) and the average bias (Avg. Bias) of channel estimates are computed via (see also, e.g., [23] and [33])

$$\text{RMSE} := \frac{1}{\|\boldsymbol{h}\|_2} \sqrt{\frac{1}{R} \sum_{r=1}^{R} \left\| \hat{\boldsymbol{h}}^{(r)} - \boldsymbol{h} \right\|_2^2}$$

$$\text{Avg. Bias} := \frac{1}{R \cdot (L+1) \|\boldsymbol{h}\|_2} \sum_{l=0}^{L} \left| \sum_{r=1}^{R} \left[\hat{\boldsymbol{h}}^{(r)}(l) - \boldsymbol{h}(l) \right] \right|$$

where

- $\hat{\boldsymbol{h}}^{(r)}$ channel estimate in the rth realization;
- R number of Monte Carlo trials;
- $\|\boldsymbol{h}\|_2$ Euclidean norm of $\boldsymbol{h}$.

The SNR is defined at the cyclostationary input of the equalizer as SNR $:= \sqrt{(\sum_{n=0}^{P-1} \mathbb{E}\{|x_1(n)|^2\}/P)/\mathbb{E}\{|v(n)|^2\}}$, where $x_1(n)$ stands for the noiseless channel output.

Experiment 1—Estimation of Frequency Offset: In order to verify the efficacy of the proposed FO estimation algorithms, we have run Monte Carlo simulations with BPSK symbols and computed the root-mean square error $[10 \log \mathbb{E}(\hat{\theta}_e - \theta_e)^2]$ of the estimator (32) and compared it with the asymptotic limit (38) and the stochastic CRB (41). In Fig. 6(a) and (b), the experimental RMSE values of FO estimator (32), which is represented by the dash-dotted line, are plotted assuming the SNR range: [0, 40] dB and two different sample sizes $T = 100$ and $T = 50$ samples, respectively. The amount of zero padding was $T_{zp} = 40\,000$. The experimental RMSE values were obtained by averaging over $R = 100$ Monte Carlo trials. The channel output was *a priori* equalized using a minimum mean square equalizer (MMSE) with 50 taps [28]. The asymptotic limit (38) is denoted by the dashed line, whereas the CRB (41) is represented by the solid line. Fig. 6(b) shows that the asymptotic limit (38) is achieved for a limited number of samples $T = 50$ and at a low SNR (10 dB). Although the proposed algorithm does not

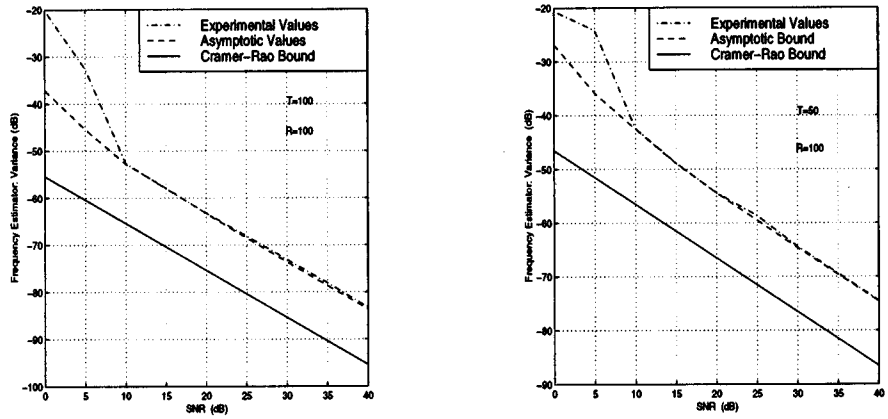

Fig. 6. Standard deviation of FO estimator (BPSK).

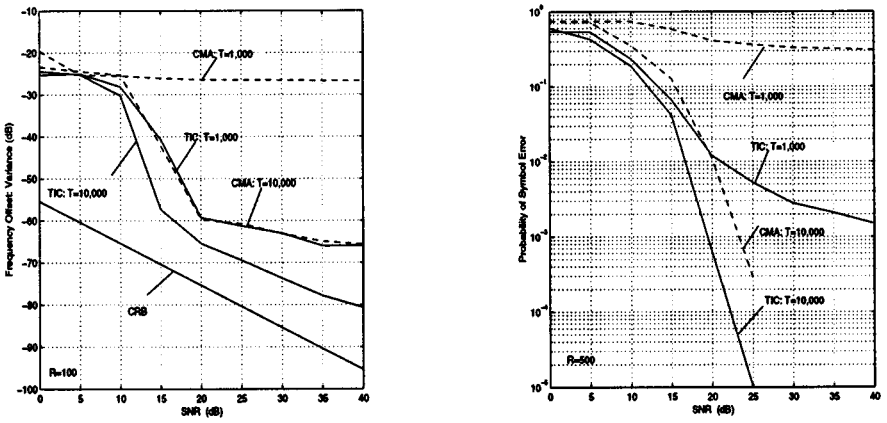

Fig. 7. TIC versus CMA: Comparisons (QPSK).

attain the stochastic CRB, Fig. 6(a) reveals the excellent performance of the FO estimator (32). For as low as $T = 100$ samples and SNR ≈ 15 dB, it achieves a variance of 10^{-6}.

Experiment 2—Comparison of TIC with CMA: In this experiment, we assess the performance of the proposed TIC set up and compare it with the CMA setup of [19] in terms of equalizing the unknown channel $h(n)$ and estimating the carrier frequency offset θ_e. Since the channel $h(n)$ has two spectral nulls, for both simulation setups, the equalizer length (50 taps) is chosen sufficiently long to compensate well the ISI effects. Input $s(n)$ is drawn from a QPSK constellation, additive noise $v(n)$ is white and normally distributed, and as before, the frequency offset $\theta_e = \pi/30$. The fourth-order moment CMA criterion of [19] is tested, assuming two different sample records available during each Monte Carlo trial: $T = 1000$ samples with adaptation step size $\mu = 0.001$ and $T = 10\,000$ samples with adaptation step size $\mu = 0.000\,34$. In both cases, the step size was optimally chosen using trial searches. For TIC setup, during each Monte Carlo trial, the unknown channel $g(n)$ is first estimated using the subspace approach, and its estimate is then plugged into the expression of the MMSE equalizer [28]. For the CMA setup, during each Monte Carlo run, a linear equalizer is estimated using the algorithm [19]. Both setups assume that frequency offset estimation is performed using (34) on the equalized channel output. The parameters of FO estimator (34) are $T = 100$ and $T_{zp} = 15\,000$. In Fig. 7(a), we plot the standard deviation of the frequency estimator (34) versus SNR corresponding to CMA and TIC setups. In Fig. 7(b), the probability of symbol error versus SNR is plotted, assuming an averaging of $R = 500$ Monte Carlo runs per SNR point. Both figures suggest that the TIC setup presents almost one order of magnitude faster convergence than the CMA. We performed quite extensive simulation experiments using the fast implementation of the block-based CMA proposed in [2] and the fractionally spaced CMA (FS-CMA), using a reduced number of samples $T = 200$ [see Fig. 10(a)]. The poor performance of the CMA algorithms calls for thorough performance and convergence analyses of the block/FS-CMA algorithms and for developing fast converging implementations before drawing any definite conclusion.

Experiment 3—Performance of Subspace Channel Estimator: In Fig. 8, the RMSE and average bias of the subspace channel estimates are plotted versus SNR in the presence and absence of FO. A number of $R = 100$ Monte Carlo runs are averaged, assuming two different sample records $T = 500$ and $T = 1000$ samples, respectively. Input $s(n)$ was drawn equiprobable from a QPSK constellation. The solid line represents the experimental values achieved by the subspace approach (23) in the case of a channel with FO, whereas the dashed line depicts the experimental values achieved by the subspace approach, assuming perfect carrier synchronization.

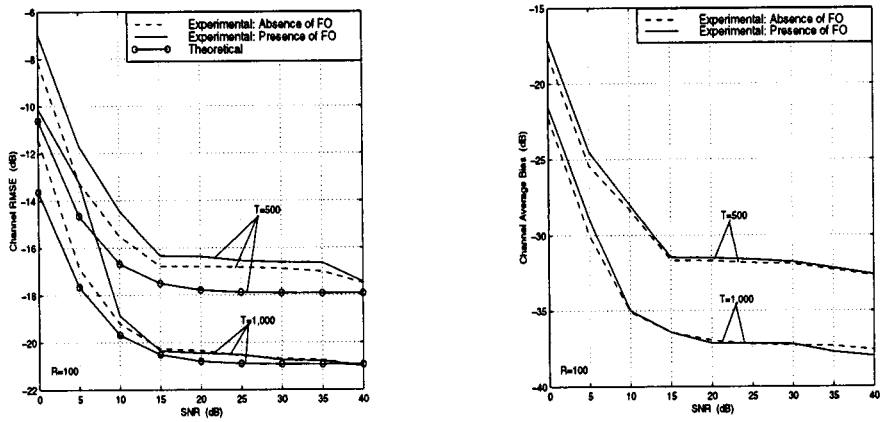

Fig. 8. Channel Avg. RMSE/bias versus SNR.

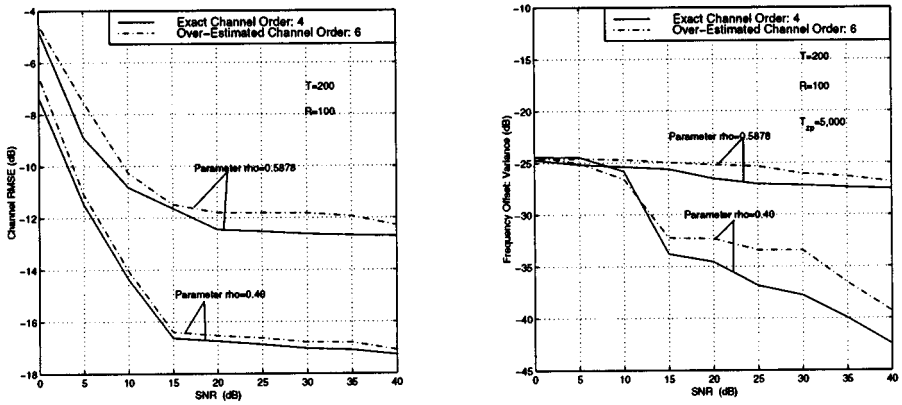

Fig. 9. Channel Avg. RMSE/frequency variance versus SNR. Exact/overestimated channel order.

In Fig. 8(a), the asymptotic variance (42) of the subspace approach is represented by the solid line with circles. Fig. 8 confirms our theoretical finding that the performance of the subspace approach in the presence of FO achieves, asymptotically, the performance of the subspace approach, assuming perfect carrier synchronization. In Fig. 9(a), the RMSE values of the subspace channel estimates are plotted versus SNR assuming $T = 200$ samples, $R = 100$ Monte Carlo runs, and two different modulating precoders ($\rho = 0.5878$ and $\rho = 0.4$) in the presence of exact $(L = 4)$ and overestimated $(L = 6)$ channel orders. In Fig. 9(b), the variance of the frequency estimator (32) is plotted versus SNR in the presence of exact and overestimated channel orders ($T_{zp} = 5000$). Fig. 9(a) and (b) shows that the subspace channel and FO estimators are quite robust to channel order overestimation errors.

Experiment 4 —Equalization Performance: In this experiment, we test if compensating for the carrier frequency offset affects the symbol error rate. In Fig. 10(b), we have plotted the probability of symbol error versus SNR for two scenarios: in the presence (solid line) and absence (dashed line) of FO, respectively. An MMSE linear equalizer with 50 taps is used for both scenarios. For each one of the $R = 500$ Monte Carlo runs, T samples are used in order to compute the subspace channel estimate. During each Monte Carlo trial, the MMSE linear equalizer of [28] is constructed using the subspace channel estimate. We have repeated this experiment for three different data records $T = 500$, 1000, and 2000 samples. The set of plots represented in Fig. 10(b) suggests that the same symbol error curves are achievable in the presence or absence of FO. From an equalization viewpoint, carrier frequency offset compensation affects performance minimally.

VIII. Conclusions

A blind channel and carrier frequency offset estimator has been proposed. Unlike existing approaches, the subspace channel estimator is resilient to order overestimation errors, color of additive stationary noise, and location of channel zeros and does not require minimization of a nonlinear criterion. A computationally simple FO estimator, whose performance is robust to residual ISI, has also been proposed. It has been shown that compensating for FO does not introduce any penalty in the asymptotic performance of the subspace channel estimation approach when compared with a perfect carrier synchronization scenario. Extension of the proposed work to the more general class of time- and frequency-selective channels described in [15] represents an important future research direction. Preliminary investigations show that the present results can be extended to compensate possibly different Doppler effects that are induced by the relative motion of the mobile.

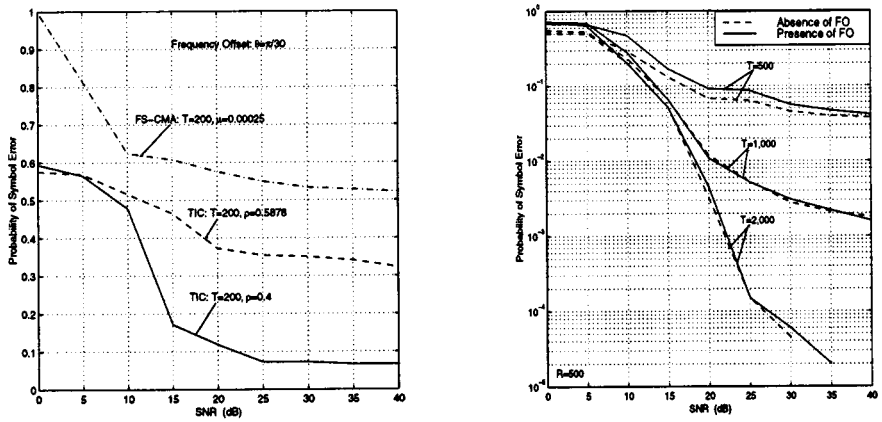

Fig. 10. Probability of symbol error.

APPENDIX A
ASYMPTOTIC ANALYSIS OF FO-ESTIMATOR

We follow the lines of proof presented in [5], [20], and [21]. Define the new variables:

$$y(n,\boldsymbol{\theta}) := \sum_{k=0}^{P-1} \alpha_k e^{j(((2\pi k/P)+2\theta_e)n+\phi_k)} \tag{53}$$

$$\boldsymbol{y}(\boldsymbol{\theta}) := [y(0,\boldsymbol{\theta})\cdots y(T-1,\boldsymbol{\theta})]^T \tag{54}$$

$$\boldsymbol{x} := \left[x_e^2(0)\cdots x_e^2(T-1)\right]^T \tag{55}$$

$$\boldsymbol{\epsilon} := [\epsilon(0)\cdots\epsilon(T-1)]^T \tag{56}$$

$$\boldsymbol{a}(n) := \frac{\partial y(n,\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \tag{57}$$

$$\boldsymbol{A} := [\boldsymbol{a}(0);\cdots;\boldsymbol{a}(T-1)]^T \tag{58}$$

$$\boldsymbol{\theta} := [\alpha_0\phi_0\cdots\alpha_{P-1}\phi_{P-1}2\theta_e]^T.$$

It is easy to check that $\boldsymbol{a}(n) = [\exp(j(2\theta_e n+\phi_0))\ j\alpha_0\exp(j(2\theta_e n+\phi_0))\cdots\exp(j((2\pi(P-1)/P+2\theta_e)n+\phi_{P-1}))j\alpha_{P-1}\exp(j((2\pi(P-1)/P+2\theta_e)n+\phi_{P-1}))jn\sum_{k=0}^{P-1}\alpha_k\exp(j((2\pi k/P+2\theta_e)n+\phi_k))]^T$. The asymptotic distribution of $\hat{\boldsymbol{\theta}}$ is performed by first establishing the asymptotic consistency[2] of $\hat{\boldsymbol{\theta}}$, namely, $\hat{\alpha}_k-\alpha_k = o_p(1)$, $\hat{\phi}_k-\phi_k = o_p(1)$, and $T(\hat{\theta}_e-\theta_e) = o_p(1)$ and then by deriving the asymptotic distribution via a Taylor series expansion of the NLS-criterion in the neighborhood of the true values. Because the consistency of the NLS estimates can be established similar to [21, pp. 113–115] or [5, Lemma and Th. 1a], we defer its presentation. Considering the Taylor series expansion of the gradient vector $\partial J(\hat{\boldsymbol{\theta}})/\partial\boldsymbol{\theta}$ around the true value $\boldsymbol{\theta}$, we obtain

$$\frac{\partial J\left(\hat{\boldsymbol{\theta}}\right)}{\partial\boldsymbol{\theta}} \cong \frac{\partial J(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} + \frac{\partial^2 J(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^2}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right). \tag{59}$$

Since $\partial J(\hat{\boldsymbol{\theta}})/\partial\boldsymbol{\theta} = \mathbf{0}$, from (59), we obtain

$$\hat{\boldsymbol{\theta}}-\boldsymbol{\theta} \cong -\left[\frac{\partial^2 J(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^2}\right]^{-1}\frac{\partial J(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} \tag{60}$$

[2]If a_n and b_n denote two arbitrary sequences of random variables, the notations $a_n = o_P(1)$ and $b_n = O_P(1)$ denote that a_n converges to zero with probability one and that b_n is bounded with probability one, respectively.

where

$$\begin{aligned}\frac{\partial J(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} &= \frac{1}{2T}\left\{[\boldsymbol{x}-\boldsymbol{y}(\boldsymbol{\theta})]^H\left(-\frac{\partial\boldsymbol{y}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)+\left(-\frac{\partial\boldsymbol{y}^H(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)\right.\\ &\quad\left.\times[\boldsymbol{x}-\boldsymbol{y}(\boldsymbol{\theta})]\right\}\\ &= \frac{1}{T}\,\mathrm{Re}\left[-\frac{\partial\boldsymbol{y}^H(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\,\boldsymbol{\epsilon}\right]\end{aligned} \tag{61}$$

$$\begin{aligned}\frac{\partial^2 J(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^2} &= \frac{1}{T}\frac{\partial}{\partial\boldsymbol{\theta}}\,\mathrm{Re}\left[-\frac{\partial\boldsymbol{y}^H(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\,\boldsymbol{\epsilon}\right]\\ &= \frac{1}{T}\,\mathrm{Re}\left[\frac{\partial\boldsymbol{y}^H(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\cdot\frac{\partial\boldsymbol{y}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^T}-\sum_{n=0}^{T-1}\frac{\partial^2\boldsymbol{y}^H(\boldsymbol{\theta},n)}{\partial\boldsymbol{\theta}^2}\,\epsilon(n)\right]\end{aligned} \tag{62}$$

and H stands for conjugate transposition. Define the matrix

$$\boldsymbol{Y}_T := \mathrm{Re}\left[\frac{\partial\boldsymbol{y}^H(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\cdot\frac{\partial\boldsymbol{y}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^T}-\sum_{n=0}^{T-1}\frac{\partial^2\boldsymbol{y}^H(\boldsymbol{\theta},n)}{\partial\boldsymbol{\theta}^2}\,\epsilon(n)\right]. \tag{63}$$

Plugging (61) and (62) back into (60) and taking into account (63), we obtain

$$\hat{\boldsymbol{\theta}}-\boldsymbol{\theta} = \boldsymbol{Y}_T^{-1}\,\mathrm{Re}\left[\frac{\partial\boldsymbol{y}^H(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\cdot\boldsymbol{\epsilon}\right]. \tag{64}$$

Define the $(2P+1)\times(2P+1)$ block diagonal matrix $\boldsymbol{K}_T := \mathrm{diag}\{[T^{1/2}, T^{1/2}\cdots T^{1/2}, T^{3/2}]\}$. From (64), we deduce that

$$\begin{aligned}&\boldsymbol{K}_T\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)\\ &\quad= \left(\boldsymbol{K}_T^{-1}\boldsymbol{Y}_T\boldsymbol{K}_T^{-1}\right)^{-1}\left[\boldsymbol{K}_T^{-1}\,\mathrm{Re}\left(\frac{\partial\boldsymbol{y}^H(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\cdot\boldsymbol{\epsilon}\right)\right].\end{aligned} \tag{65}$$

In Appendix B, it is shown that the second term in the expansion of $\boldsymbol{K}_T^{-1}\boldsymbol{Y}_T\boldsymbol{K}_T^{-1}$ [see (63)] is asymptotically negligible as $T\to\infty$, i.e., $\boldsymbol{K}_T^{-1}\sum_{n=0}^{T-1}((\partial^2\boldsymbol{y}^H(\boldsymbol{\theta},n))/\partial\boldsymbol{\theta}^2)\epsilon(n)\boldsymbol{K}_T^{-1} = o(1)$.

From (65), we obtain

$$\begin{aligned}\lim_{T\to\infty} \boldsymbol{K}_T\mathbb{E}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)^T\boldsymbol{K}_T &= \lim_{T\to\infty}\left[\boldsymbol{K}_T^{-1}\,\mathrm{Re}\left(\frac{\partial \boldsymbol{y}^H(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\cdot\frac{\partial\boldsymbol{y}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^T}\right)\boldsymbol{K}_T^{-1}\right]^{-1}\\ &\times\left[\boldsymbol{K}_T^{-1}\mathbb{E}\,\mathrm{Re}\left(\frac{\partial\boldsymbol{y}^H(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\boldsymbol{\epsilon}\right)\mathrm{Re}\left(\frac{\partial\boldsymbol{y}^H(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\boldsymbol{\epsilon}\right)^T\boldsymbol{K}_T^{-1}\right]\\ &\times\left[\boldsymbol{K}_T^{-1}\,\mathrm{Re}\left(\frac{\partial\boldsymbol{y}^H(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\cdot\frac{\partial\boldsymbol{y}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^T}\right)\boldsymbol{K}_T^{-1}\right]. \qquad (66)\end{aligned}$$

Since $\boldsymbol{A}^H=\partial\boldsymbol{y}^H(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$, (66) can be expressed equivalently as

$$\begin{aligned}\lim_{T\to\infty}\boldsymbol{K}_T\mathbb{E}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)^T\boldsymbol{K}_T &= \lim_{T\to\infty}\left[\boldsymbol{K}_T^{-1}\,\mathrm{Re}\left(\boldsymbol{A}^H\boldsymbol{A}\right)\boldsymbol{K}_T^{-1}\right]^{-1}\\ &\times\left[\boldsymbol{K}_T^{-1}\mathbb{E}\,\mathrm{Re}\left(\boldsymbol{A}^H\boldsymbol{\epsilon}\right)\mathrm{Re}\left(\boldsymbol{A}^H\boldsymbol{\epsilon}\right)^T\boldsymbol{K}_T^{-1}\right]\\ &\times\left[\boldsymbol{K}_T^{-1}\,\mathrm{Re}\left(\boldsymbol{A}^H\boldsymbol{A}\right)\boldsymbol{K}_T^{-1}\right]^{-1}. \qquad (67)\end{aligned}$$

To compute the individual factors in the right-hand side of (67), we appeal to the following result [21].

Proposition 2: With k denoting a positive integer and $\delta(\omega)$ denoting Kronecker delta, it holds that

$$\lim_{T\to\infty}\frac{\sum_{n=0}^{T-1}n^k e^{j(\omega n+\phi)}}{T^{k+1}}=\frac{e^{j\phi}\delta(\omega)}{k+1}. \qquad (68)$$

Using (68), some direct computations show that the symmetric matrix $\lim_{T\to\infty}\boldsymbol{K}_T^{-1}\,\mathrm{Re}(\boldsymbol{A}^H\boldsymbol{A})\boldsymbol{K}_T^{-1}$ is given by

$$\lim_{T\to\infty}\boldsymbol{K}_T^{-1}\,\mathrm{Re}\left(\boldsymbol{A}^H\boldsymbol{A}\right)\boldsymbol{K}_T^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0\\ 0 & \alpha_0^2 & 0 & 0 & \cdots & 0 & 0 & \frac{\alpha_0^2}{2}\\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0\\ 0 & 0 & 0 & \alpha_1^2 & \ddots & 0 & 0 & \frac{\alpha_1^2}{2}\\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \cdots & \vdots\\ 0 & 0 & 0 & 0 & \cdots & 0 & \alpha_{P-1}^2 & \frac{\alpha_{P-1}^2}{2}\\ 0 & \frac{\alpha_0^2}{2} & 0 & \frac{\alpha_1^2}{2} & \cdots & 0 & \frac{\alpha_{P-1}^2}{2} & \frac{\sum_{k=0}^{P-1}\alpha_k^2}{3}\end{bmatrix}. \qquad (69)$$

To compute the second factor in the right-hand side of (67), the following identity is used:

$$\begin{aligned}&\mathbb{E}\,\mathrm{Re}\left(\boldsymbol{A}^H\boldsymbol{\epsilon}\right)\mathrm{Re}\left(\boldsymbol{A}^H\boldsymbol{\epsilon}\right)^T\\ &\quad=\tfrac{1}{2}\,\mathrm{Re}\left[\mathbb{E}\left(\boldsymbol{A}^H\boldsymbol{\epsilon}\boldsymbol{\epsilon}^H\boldsymbol{A}\right)\right]+\tfrac{1}{2}\,\mathrm{Re}\left[\mathbb{E}\left(\boldsymbol{A}^H\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T\boldsymbol{A}^*\right)\right]. \qquad (70)\end{aligned}$$

In Appendix C, a generic entry of the matrix $\lim_{T\to\infty}\boldsymbol{K}_T^{-1}\mathbb{E}(\boldsymbol{A}^H\boldsymbol{\epsilon}\boldsymbol{\epsilon}^H\boldsymbol{A})\boldsymbol{K}_T^{-1}$ in (70) is computed explicitly. Using similar arguments, it can be shown that $\lim_{T\to\infty}\boldsymbol{K}_T^{-1}\mathbb{E}(\boldsymbol{A}^H\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T\boldsymbol{A}^*)\boldsymbol{K}_T^{-1}=\boldsymbol{0}$.

The asymptotic covariance of the FO estimator $\lim_{T\to\infty}T^3(2\hat{\theta}_e-2\theta_e)^2$ can now be calculated and is given by the $(2P+1,\,2P+1)$ entry of the right-hand side matrix in (67). The computations are eased by noting that a closed-form expression for the inverse of matrix (69) can be simply obtained by using the block-matrix inversion formula [24, p. 413, (a.14)]. After some straightforward calculations (whose details do not present any interest), (38) is obtained for the FO estimator's asymptotic variance. □

Appendix B
Proof of $\lim_{T\to\infty}\boldsymbol{K}_T^{-1}\sum_{n=0}^{T-1}(\partial^2\boldsymbol{y}^*(\boldsymbol{\theta},n)/\partial\boldsymbol{\theta}^2)\boldsymbol{\epsilon}(n)\boldsymbol{K}_T^{-1}=o(1)$

A generic entry of the matrix $\boldsymbol{K}_T^{-1}\sum_{n=0}^{T-1}(\partial^2\boldsymbol{y}^*(\boldsymbol{\theta},n)/\partial\boldsymbol{\theta}^2)\boldsymbol{\epsilon}(n)\boldsymbol{K}_T^{-1}$ has the form

$$q(n):=\frac{\sum_{n=0}^{T-1}n^l e^{-j(((2\pi k/P)+2\theta_e)n+\phi_k)}\epsilon(n)}{T^{l+1}} \qquad (71)$$

where $l=0,\,1,\,2$, and $k=0,\cdots,P-1$. Since $v(n)$ satisfies the mixing condition (35), $s(n)$ has finite moments, and $h(n)$ has finite memory, it follows that $\epsilon(n)$ [which is defined in (37)] satisfies the mixing condition

$$\sum_{\tau=-\infty}^{\infty}|c_{\epsilon\epsilon}(k;\tau)|<\infty,\quad\forall\,k=0,1,\cdots,P-1 \qquad (72)$$

where $c_{\epsilon\epsilon}(n;\tau):=\mathbb{E}\epsilon^*(n)\epsilon(n+\tau)$. We show next that $\lim_{T\to\infty}\mathbb{E}|q(n)|^2=0$. Straightforward computations show that

$$\mathbb{E}|q(n)|^2=\frac{\sum_{m=0}^{T-1}\sum_{\tau=-m}^{T-1-m}m^l(m+\tau)^l e^{-j((2\pi k/P)+2\theta_e)\tau}c_{\epsilon\epsilon}(m;\tau)}{T^{2l+2}}. \qquad (73)$$

Using the inequality $m^l(m+\tau)^l\le T^{2l},\,\forall m=0,\cdots,T-1$, we deduce from (73) the following relations:

$$\begin{aligned}\mathbb{E}|q(n)|^2 &\le\frac{\sum_{m=0}^{T-1}\sum_{\tau=-m}^{T-1-m}m^l(m+\tau)^l\,|c_{\epsilon\epsilon}(m;\tau)|}{T^{2l+2}}\\ &\le\frac{T^{2l}\sum_{m=0}^{T-1}\sum_{\tau=-T+1}^{T-1}|c_{\epsilon\epsilon}(m;\tau)|}{T^{2l+2}}\\ &=\frac{1}{TP}\sum_{k=0}^{P-1}\sum_{\tau=-T+1}^{T-1}|c_{\epsilon\epsilon}(k;\tau)|\overset{T\to\infty}{\to}0\end{aligned}$$

which proves the assertion. □

APPENDIX C
CALCULATION OF $\lim_{T\to\infty} \boldsymbol{K}_T^{-1}\mathbb{E}(\boldsymbol{A}^H\boldsymbol{\epsilon}\boldsymbol{\epsilon}^H\boldsymbol{A})\boldsymbol{K}_T^{-1}$

Since a generic entry of the vector $\boldsymbol{A}^H\boldsymbol{\epsilon}$ is of the form $\sum_{n=0}^{T-1} n^l\alpha_k \exp(-j((2\pi k/P + 2\theta_e)n + \phi_k))\epsilon(n)$, with $l = 0, 1$, and $k = 0, \cdots, P-1$, it follows that the calculation of an arbitrary entry of the matrix $\lim_{T\to\infty} \boldsymbol{K}_T^{-1}\mathbb{E}(\boldsymbol{A}^H\boldsymbol{\epsilon}\boldsymbol{\epsilon}^H\boldsymbol{A})\boldsymbol{K}_T^{-1}$ reduces to the evaluation of the limit

$$\gamma := \lim_{T\to\infty} \frac{\alpha_{k_0}\alpha_{k_1}e^{-j(\phi_{k_0}-\phi_{k_1})}}{T^{l_0+l_1+1}} \times \left\{ \sum_{n_0=0}^{T-1}\sum_{n_1=0}^{T-1} n_0^{l_0} n_1^{l_1}\mathbb{E}\epsilon(n_0)\epsilon^*(n_1) \times e^{-j((2\pi k_0/P)+2\theta_e)n_0} e^{j((2\pi k_1+2\theta_e/P))n_1} \right\} \quad (74)$$

where k_0, $k_1 = 0, \cdots, P-1$, and l_0, $l_1 = 0, 1$. Considering the change of variables $n_1 \leftrightarrow n$ and $n_0 \leftrightarrow n+\tau$, (74) can be expressed equivalently as

$$\gamma = \frac{\beta}{T^{l_0+l_1+1}} \sum_{n=0}^{T-1}\sum_{\tau=-n}^{T-1-n} n^{l_1}(n+\tau)^{l_0} c_{\epsilon\epsilon}(n;\tau) \times e^{(j2\pi(k_1-k_0)n/P)} e^{-j((2\pi k_0/P)+2\theta_e)\tau} \quad (75)$$

where $\beta := \alpha_{k_0}\alpha_{k_1}\exp(-j(\phi_{k_0}-\phi_{k_1}))$. Next, the right-hand side double sum in (75) is decomposed into three sums, and each sum will be evaluated individually:

$$\gamma = \beta \lim_{T\to\infty} \left\{ \frac{1}{T^{l_0+l_1+1}} \sum_{n=0}^{T-1}\sum_{\tau=-(T-1)}^{T-1} (\cdot) - \frac{1}{T^{l_0+l_1+1}} \times \sum_{n=0}^{T-1}\sum_{\tau=T-n}^{T-1} (\cdot) - \frac{1}{T^{l_0+l_1+1}} \sum_{n=0}^{T-1}\sum_{\tau=-(T-1)}^{-n-1} (\cdot) \right\}. \quad (76)$$

Denote the first term in the right-hand side of (76) by γ_1. By interchanging the order of summation, γ_1 can be further expressed as

$$\gamma_1 = \beta \lim_{T\to\infty} \sum_{\tau=-T+1}^{T-1} e^{-j((2\pi k_0/P)+2\theta_e)\tau} \times \left\{ \frac{1}{T^{l_0+l_1+1}} \sum_{n=0}^{T-1} n^{l_1}(n+\tau)^{l_0} c_{\epsilon\epsilon}(n;\tau) \times e^{(j2\pi(k_1-k_0)n/P)} \right\}. \quad (77)$$

Let us now suppose that τ is fixed, and evaluate the inner term of (77)

$$\lim_{T\to\infty} \frac{1}{T^{l_0+l_1+1}} \sum_{n=0}^{T-1} n^{l_1}(n+\tau)^{l_0} c_{\epsilon\epsilon}(n;\tau) e^{(j2\pi(k_1-k_0)n/P)}. \quad (78)$$

Using Proposition 2 and the mixing condition (35), it is easy to check that

$$\lim_{T\to\infty} \frac{1}{T^{l_0+l_1+1}} \sum_{n=0}^{T-1} n^l c_{\epsilon\epsilon}(n;\tau) e^{(j2\pi(k_1-k_0)n/P)} = 0 \quad \forall l < l_0+l_1. \quad (79)$$

Considering the binomial expansion of $(n+\tau)^{l_0}$ in (78) and keeping in mind (79), it follows that the limit (78) reduces to the calculation of

$$\lim_{T\to\infty} \frac{1}{T^{l_0+l_1+1}} \sum_{n=0}^{T-1} n^{l_0+l_1} c_{\epsilon\epsilon}(n;\tau) e^{(j2\pi(k_1-k_0)n/P)}. \quad (80)$$

Considering the change of variables $n = pP+r, 0 \le r \le P-1$, and $p = 0, \cdots, \lceil((T-1)/P)\rceil$, (80) can be expressed as

$$\lim_{T\to\infty} \frac{1}{T^{l_0+l_1+1}} \sum_{p=0}^{\lceil((T-1)/P)\rceil} \sum_{r=0}^{P-1} (pP+r)^{l_0+l_1} c_{\epsilon\epsilon}(r;\tau) \times e^{(j2\pi(k_1-k_0)r/P)}. \quad (81)$$

Plugging the binomial expansion of $(pP+r)^{l_0+l_1}$ into (81), taking into account (79), and interchanging the order of summation, (81) reduces to

$$\begin{aligned} &\sum_{r=0}^{P-1} c_{\epsilon\epsilon}(r;\tau) e^{(j2\pi(k_1-k_0)r/P)} \times \left\{ \lim_{T\to\infty} \sum_{p=0}^{\lceil((T-1)/P)\rceil} (pP+r)^{l_0+l_1} \right\} \\ &= \frac{1}{P(l_0+l_1+1)} \sum_{r=0}^{P-1} c_{\epsilon\epsilon}(r;\tau) e^{(j2\pi(k_1-k_0)r/P)} \\ &= \frac{1}{l_0+l_1+1} C_{\epsilon\epsilon}(k_0-k_1;\tau) \end{aligned} \quad (82)$$

where $C_{\epsilon\epsilon}(k_0-k_1;\tau)$ stands for the cyclic correlation coefficient [see (5)]. Plugging (82) back into (77), it follows that

$$\begin{aligned} \gamma_1 &= \beta \lim_{T\to\infty} \sum_{\tau=-T+1}^{T-1} \frac{1}{l_0+l_1+1} C_{\epsilon\epsilon}(k_0-k_1;\tau) \times e^{-j((2\pi k_0/P)+2\theta_e)\tau} \\ &= \frac{\beta}{l_0+l_1+1} S_{\epsilon\epsilon}\left(k_0-k_1; \frac{2\pi k_0}{P}+2\theta_e\right). \end{aligned} \quad (83)$$

Next, we show that the second term on the right-hand side of (76) is equal to zero. Denote this term by γ_2

$$\gamma_2 := \lim_{T\to\infty} \frac{\beta}{T^{l_0+l_1+1}} \sum_{n=0}^{T-1}\sum_{\tau=T-n}^{T-1} n^{l_1}(n+\tau)^{l_0} c_{\epsilon\epsilon}(n;\tau) \times e^{(j2\pi(k_1-k_0)n/P)} e^{-j((2\pi k_0/P)+2\theta_e)\tau}. \quad (84)$$

Due to the mixing condition (35), for any arbitrary $\delta > 0$, there exists an integer $N(\delta)$ such that

$$\sum_{\tau>N(\delta)} |c_{\epsilon\epsilon}(n;\tau)| < \delta, \quad \forall n = 0, \cdots, P-1. \quad (85)$$

We split the sum (84) in two terms

$$\gamma_2 = \lim_{T\to\infty}\left\{\frac{\beta}{T^{l_0+l_1+1}}\sum_{n=0}^{T-N(\delta)-1}\sum_{\tau=T-n}^{T-1}(\cdot) - \frac{\beta}{T^{l_0+l_1+1}}\sum_{n=T-N(\delta)}^{T-1}\sum_{\tau=T-n}^{T-1}(\cdot)\right\}. \quad (86)$$

We next show that both terms on the right-hand side of (86) converge to zero as $T\to\infty$. Since $n^{l_1}(n+\tau)^{l_0} \le T^{l_0+l_1}$ for $n=0,\cdots,T-1$ and $\tau = T-n$, we can upper bound the first term on the right-hand side of (86) as follows:

$$\begin{aligned}
&\frac{\beta}{T^{l_0+l_1+1}}\sum_{n=0}^{T-N(\delta)-1}\sum_{\tau=T-n}^{T-1} n^{l_1}(n+\tau)^{l_0}c_{\epsilon\epsilon}(n;\tau)\\
&\quad\times e^{(j2\pi(k_1-k_0)n/P)}e^{-j((2\pi k_0/P)+2\theta_e)\tau}\\
&\le \frac{|\beta|}{T^{l_0+l_1+1}}\sum_{n=0}^{T-1}\sum_{\tau=N(\delta)+1}^{T-1} T^{l_0+l_1}\,|c_{\epsilon\epsilon}(n;\tau)|\\
&\le \frac{|\beta|}{P}\sum_{r=0}^{P-1}\sum_{\tau=N(\delta)+1}^{\infty}|c_{\epsilon\epsilon}(r;\tau)|\\
&\le |\beta|\delta \overset{\delta\to 0}{=} 0 \qquad (87)
\end{aligned}$$

where in deriving the last two inequalities in (87), we used the change of variables $n = pP + r$, $r = 0,\cdots,P-1$, and (85). Similarly, we can upper bound the second term on the right-hand side of (86) as

$$\begin{aligned}
&\left|\frac{\beta}{T^{l_0+l_1+1}}\sum_{n=T-N(\delta)}^{T-1}\sum_{\tau=T-n}^{T-1} n^{l_1}(n+\tau)^{l_0}c_{\epsilon\epsilon}(n;\tau)\right.\\
&\quad\left.\times\, e^{(j2\pi(k_1-k_0)n/P)}e^{-j((2\pi k_0/P)+2\theta_e)\tau}\right|\\
&\le \frac{|\beta|}{T^{l_0+l_1+1}}\sum_{n=T-N(\delta)}^{T-1}\sum_{\tau=1}^{T-1} T^{l_0+l_1}\,|c_{\epsilon\epsilon}(n;\tau)|\\
&\le \frac{|\beta|N(\delta)}{T}\max_{n=0,\cdots,P-1}\left\{\sum_{\tau=1}^{T-1}|c_{\epsilon\epsilon}(n;\tau)|\right\}\\
&\overset{T\to\infty}{=} 0. \qquad (88)
\end{aligned}$$

Thus, $\lim_{T\to\infty}\gamma_2 = 0$. In a similar way, the third term on the right-hand side of (76) is shown to converge to zero. This concludes the proof. □

Appendix D
Proof (Sketch) of Theorem 1

The LS-solution of (23) is given by

$$\hat{\boldsymbol{h}}_1 = \left(\hat{\boldsymbol{A}}_1^H\hat{\boldsymbol{A}}_1\right)^{-1}\left(\hat{\boldsymbol{A}}_1^H\hat{\boldsymbol{b}}_1\right). \quad (89)$$

Since $\boldsymbol{h}_1 = (\hat{\boldsymbol{A}}_1^H\hat{\boldsymbol{A}}_1)^{-1}(\hat{\boldsymbol{A}}_1^H\hat{\boldsymbol{A}}_1\boldsymbol{h}_1)$, from (89), it follows that

$$\hat{\boldsymbol{h}}_1 - \boldsymbol{h}_1 = \left(\hat{\boldsymbol{A}}_1^H\hat{\boldsymbol{A}}_1\right)^{-1}\hat{\boldsymbol{A}}_1^H\left(\hat{\boldsymbol{b}}_1 - \hat{\boldsymbol{A}}_1\boldsymbol{h}_1\right). \quad (90)$$

Since $\boldsymbol{b}_1 - \boldsymbol{A}_1\boldsymbol{h}_1 = \boldsymbol{0}$, (90) can be expressed as

$$\hat{\boldsymbol{h}}_1 - \boldsymbol{h}_1 = \left(\hat{\boldsymbol{A}}_1^H\hat{\boldsymbol{A}}_1\right)^{-1}\hat{\boldsymbol{A}}_1^H\left[\left(\hat{\boldsymbol{b}}_1 - \boldsymbol{b}_1\right) - \left(\hat{\boldsymbol{A}}_1 - \boldsymbol{A}_1\right)\boldsymbol{h}_1\right]. \quad (91)$$

Because matrix $(\hat{\boldsymbol{b}}_1-\boldsymbol{b}_1)-(\hat{\boldsymbol{A}}_1-\boldsymbol{A}_1)\boldsymbol{h}_1$ depends linearly on the entries of the vector $\hat{\boldsymbol{c}}_{xx}-\boldsymbol{c}_{xx}$, there exists a selection matrix $\boldsymbol{D}$ whose entries are not dependent on $\hat{\boldsymbol{c}}_{xx}-\boldsymbol{c}_{xx}$ such that

$$\left(\hat{\boldsymbol{b}}_1 - \boldsymbol{b}_1\right) - \left(\hat{\boldsymbol{A}}_1 - \boldsymbol{A}_1\right)\boldsymbol{h}_1 = \boldsymbol{D}\left(\hat{\boldsymbol{c}}_{xx} - \boldsymbol{c}_{xx}\right). \quad (92)$$

Simple calculations show that $\boldsymbol{D}$ is indeed given by (43). Inserting (92) in (91), we obtain

$$\sqrt{T}\left(\hat{\boldsymbol{h}}_1 - \boldsymbol{h}_1\right) = \left(\hat{\boldsymbol{A}}_1^H\hat{\boldsymbol{A}}_1\right)^{-1}\hat{\boldsymbol{A}}_1^H\boldsymbol{D}\sqrt{T}\left(\hat{\boldsymbol{c}}_{xx} - \boldsymbol{c}_{xx}\right). \quad (93)$$

It has been shown that $\sqrt{T}(\hat{\boldsymbol{c}}_{xx}-\boldsymbol{c}_{xx})$ has a limiting zero mean Gaussian distribution with covariance matrix $\boldsymbol{\Sigma}$ [12]. It follows that $\sqrt{T}(\hat{\boldsymbol{c}}_{xx}-\boldsymbol{c}_{xx}) = O(1)$. However, if $\hat{\theta}_e = \theta_e + O(T^{-3/2})$, then for an arbitrary continuous function $\psi(\theta_e)$, we have $\psi(\hat{\theta}_e) = \psi(\theta_e) + o(T^{-3/2})$ [24, th. C1,2, pp. 422 and 423]. It follows that asymptotically, the estimates $\hat{\theta}_e$ and $\hat{\boldsymbol{A}}_1$ can be substituted by their true values in (93) without affecting the asymptotic distribution of $\sqrt{T}(\hat{\boldsymbol{h}}_1 - h_1)$. From (93), the asymptotic covariance (42) is obtained. □

Appendix E
Proof of Proposition 3

Since

$$\begin{aligned}
&T\operatorname{cov}\left(\hat{C}_{xx}(k_1,\tau_1),\hat{C}_{xx}(k_2,\tau_2)\right)\\
&\quad= T\mathbb{E}\left[\left(\hat{C}_{xx}(k_1,\tau_1) - C_{xx}(k_1,\tau_1)\right)\right.\\
&\qquad\left.\times\left(\hat{C}_{xx}(k_2,\tau_2) - C_{xx}(k_2,\tau_2)\right)^*\right] \qquad (94)
\end{aligned}$$

it follows that the computation of the asymptotic covariance (44) resumes to computing

$$\begin{aligned}
&\mathbb{E}\left[\hat{C}_{xx}(k_1,\tau_1)\,\hat{C}_{xx}^*(k_2,\tau_2)\right]\\
&\quad\overset{c.f.(20)}{=}\frac{1}{T^2}\sum_{n_1,n_2=0}^{T-1}\\
&\qquad\times\mathbb{E}\left[x(n_1+\tau_1)x^*(n_1)x^*(n_2+\tau_2)x(n_2)\right]\\
&\qquad\times e^{(-j2\pi(k_1n_1-k_2n_2)/P)}.
\end{aligned}$$

The following decomposition holds

$$\begin{aligned}
&\mathbb{E}\left[x(n_1+\tau_1)x^*(n_1)x^*(n_2+\tau_2)x(n_2)\right]\\
&\quad=\mathbb{E}\left[x(n_1+\tau_1)x^*(n_1)\right]\mathbb{E}\left[x^*(n_2+\tau_2)x(n_2)\right]\\
&\qquad+\mathbb{E}\left[x(n_1+\tau_1)x^*(n_2+\tau_2)\right]\mathbb{E}\left[x^*(n_1)x(n_2)\right]\\
&\qquad+\mathbb{E}\left[x(n_1+\tau_1)x(n_2)\right]\mathbb{E}\left[x^*(n_1)x^*(n_2+\tau_2)\right]\\
&\qquad+\operatorname{cum}\left(x(n_1+\tau_1),x^*(n_1),x^*(n_2+\tau_2),x(n_2)\right). \qquad (95)
\end{aligned}$$

The contribution of the first term on the right-hand side term of (95) corresponds to the product of means, and it cancels out, considering all the terms in the entire decomposition on

the right-hand side of (94). The third term on the right-hand side of (95) is zero due to the circularity assumption on $s(n)$ ($\mathbb{E}s(n)s(n+\tau)=0, \forall \tau, n$). Next, we compute the contribution of the second term on the right-hand side of (95). We have the following relations:

$$\begin{aligned}\zeta_1 :=& \frac{1}{T}\sum_{n_1,n_2=0}^{T-1}\mathbb{E}\left[x\left(n_1+\tau_1\right)x^*\left(n_2+\tau_2\right)\right]\\ &\times\mathbb{E}\left[x^*\left(n_1\right)x\left(n_2\right)\right]e^{(-j2\pi(k_1n_1-k_2n_2)/P)}\\ =&\frac{1}{T}\sum_{n_1,n_2=0}^{T-1}c_{xx}\left(n_2+\tau_2;\, n_1-n_2+\tau_1-\tau_2\right)\\ &\times c^*_{xx}\left(n_2;\, n_1-n_2\right)e^{(-j2\pi(k_1n_1-k_2n_2)/P)}. \end{aligned}\tag{96}$$

Plugging

$$c_{xx}(n;\,\tau)=\sum_{k=0}^{P-1}C_{xx}(k;\,\tau)e^{(j2\pi nk/P)}$$

into (96), we can express ζ_1 as follows:

$$\begin{aligned}\zeta_1=&\sum_{l_1,l_2=0}^{P-1}\frac{1}{T}\sum_{n_1,n_2=0}^{T-1}C_{xx}\left(l_1;\, n_1-n_2+\tau_1-\tau_2\right)\\ &\times e^{(j2\pi l_1(n_2+\tau_2)/P)}C^*_{xx}\left(l_2;\, n_1-n_2\right)\\ &\times e^{(-j2\pi l_2n_2/P)}e^{(-j2\pi k_1n_1/P)}e^{j(2\pi k_2n_2/P)}. \end{aligned}\tag{97}$$

Considering the change of variables $m_1=n_1-n_2$ and $m_2=n_2$, we can decompose the right-hand side of (97) into the sum

$$\zeta_1=B_1+C_1\tag{98}$$

where

$$\begin{aligned}B_1:=&\sum_{m_1=0}^{T-1}\sum_{m_2=0}^{T-1-m_1}\frac{e^{(j2\pi l_1\tau_2/P)}}{T}C_{xx}\left(l_1;\, m_1+\tau_1-\tau_2\right)\\ &\times C^*_{xx}\left(l_2;\, m_1\right)e^{(-j2\pi k_1m_1/P)}e^{(j2\pi(l_1-l_2+k_2-k_1)m_2/P)}\end{aligned}\tag{99}$$

$$\begin{aligned}C_1:=&\sum_{m_1=-T+1}^{-1}\sum_{m_2=-m_1}^{T-1}\frac{e^{(j2\pi l_1\tau_2/P)}}{T}C_{xx}\left(l_1;\, m_1+\tau_1-\tau_2\right)\\ &\times C^*_{xx}\left(l_2;\, m_1\right)e^{(-j2\pi k_1m_1/P)}e^{(j2\pi(l_1-l_2+k_2-k_1)m_2/P)}.\end{aligned}\tag{100}$$

Next, we compute the term B_1. We distinguish two cases. If $l_1-l_2+k_2-k_1=0$, we then have

$$\begin{aligned}B_1=&\frac{e^{(j2\pi l_1\tau_2/P)}}{T}\sum_{m_1=0}^{T-1}C_{xx}\left(l_1;\, m_1+\tau_1-\tau_2\right)C^*_{xx}\left(l_2;\, m_1\right)\\ &\times e^{(-j2\pi k_1m_1/P)}\left(T-m_1\right)\end{aligned}$$

and using well-known results on Césaro sums, we obtain that

$$\begin{aligned}\lim_{T\to\infty}B_1=&e^{(j2\pi l_1\tau_2/P)}\sum_{m_1=0}^{\infty}C_{xx}\left(l_1;\, m_1+\tau_1-\tau_2\right)\\ &\times C^*_{xx}\left(l_2;\, m_1\right)e^{(-j2\pi k_1m_1/P)}.\end{aligned}\tag{101}$$

If $l_1-l_2+k_2-k_1\neq 0$ since $\sum_{m_2=0}^{T-1-m_1}\exp(j2\pi(l_1-l_2+k_2-k_1)m_2/P)$ is uniformly bounded in T, and $\sum_{m_1}|C_{xx}(l_1;\, m_1+\tau_1-\tau_2)C_{xx}(l_2;\, m_1)|<\infty$, we have that $\lim_{T\to\infty}B_1=0$. Similar computations lead to

$$\begin{aligned}\lim_{T\to\infty}C_1=&e^{(j2\pi l_1\tau_2/P)}\sum_{m_1=-\infty}^{-1}C_{xx}\left(l_1;\, m_1+\tau_1-\tau_2\right)\\ &\times C^*_{xx}\left(l_2;\, m_1\right)e^{(-j2\pi k_1m_1/P)}.\end{aligned}\tag{102}$$

From (101) and (102), we obtain that

$$\begin{aligned}\lim_{T\to\infty}\zeta_1=&\sum_{\substack{l_1,l_2=0\\ l_1-l_2=k_1-k_2}}^{P-1}\left(\sum_{m_1}C_{xx}\left(l_1;\, m_1+\tau_1-\tau_2\right)\right.\\ &\left.\times\, C^*_{xx}\left(l_2;\, m_1\right)e^{(-j2\pi k_1m_1/P)}\right)e^{(j2\pi l_1\tau_2/P)}\end{aligned}\tag{103}$$

and using the Parserval formula, the right-hand side of (103) reduces to ξ_1 [see (45)] [6].

Next, we compute the fourth term on the right-hand side of (95), i.e., the cumulant term, and show it to be equal asymptotically with ξ_2 [see (46)]. Define

$$\begin{aligned}\zeta_2:=&\frac{1}{T}\operatorname{cum}\left(x\left(n_1+\tau_1\right),\, x^*\left(n_1\right),\, x^*\left(n_2+\tau_2\right),x\left(n_2\right)\right)\\ &\times e^{(-j2\pi(k_1n_1-k_2n_2)/P)}.\end{aligned}\tag{104}$$

Since $v(n)$ is white and independently distributed of $s(n)$, we have

$$\begin{aligned}&\operatorname{cum}\left(x\left(n_1+\tau_1\right),\, x^*\left(n_1\right),\, x^*\left(n_2+\tau_2\right),\, x\left(n_2\right)\right)\\ &\quad=\operatorname{cum}\left(x_1\left(n_1+\tau_1\right),\, x_1^*\left(n_1\right),\, x_1^*\left(n_2+\tau_2\right),\, x_1\left(n_2\right)\right).\end{aligned}$$

Since $f(n)$ is a periodic sequence, we can consider the Fourier series decomposition

$$f(n)=\sum_{p=0}^{P-1}F_1(p)e^{(j2\pi pn/P)}.$$

Due to the linearity of cumulants, we further express $\operatorname{cum}(x_1(n_1+\tau_1),\, x_1^*(n_1),\, x_1^*(n_2+\tau_2),\, x_1(n_2))$ as

$$\begin{aligned}&\operatorname{cum}\left(x_1\left(n_1+\tau_1\right),\, x_1^*\left(n_1\right),\, x_1^*\left(n_2+\tau_2\right),\, x_1\left(n_2\right)\right)\\ &\quad=\sum_{p_1,p_2,p_3,p_4=0}^{P-1}F_1\left(p_1\right)F_1^*\left(p_2\right)F_1^*\left(p_3\right)F_4\left(p_4\right)\\ &\qquad\times e^{(j2\pi p_1(n_1+\tau_1-l_1)/P)}e^{(-j2\pi p_2(n_1-l_2)/P)}\\ &\qquad\times e^{(-j2\pi p_3(n_2+\tau_2-l_3)/P)}e^{(j2\pi p_4(n_2-l_4)/P)}\\ &\qquad\times\sum_{l_1,l_2,l_3,l_4=0}^{L}h\left(l_1\right)h^*\left(l_2\right)h^*\left(l_3\right)h\left(l_4\right)\\ &\qquad\operatorname{cum}\left(s\left(n_1+\tau_1-l_1\right)s^*\left(n_1-l_2\right),\right.\\ &\qquad\left.s^*\left(n_2+\tau_2-l_3\right),\, s\left(n_2-l_4\right)\right).\end{aligned}\tag{105}$$

The cumulant term $\operatorname{cum}(s(n_1+\tau_1-l_1),\, s^*(n_1-l_2),\, s^*(n_2+\tau_2-l_3),\, s(n_2-l_4))$ in (105) is nonzero only if its indices are equal, i.e., $n_2-n_1=l_4-l_2$, $l_1=l_2+\tau_1$, and $l_3=l_4+\tau_2$. It follows easily that the general term on the right-hand side of

(104) is equal to

$$\begin{aligned}\kappa :=& F_1(p_1)F_1^*(p_2)F_1^*(p_3)F_1(p_4)h(l_1)h^*(l_2)h^*(l_3)h(l_4)\\ &\times e^{(j2\pi(\tau_1 p_1 - l_1 p_1 + l_2 p_2 + l_3 p_3 - \tau_2 p_3 - l_4 p_4)/P)}\\ &\times \gamma_{4s}\frac{1}{T}\sum_{n_2=n_1+l_4-l_2} e^{(j2\pi(p_1-p_2-k_1)n_1/P)}\\ &\times e^{(j2\pi(p_4-p_3+k_2)n_2/P)}\end{aligned} \tag{106}$$

where $\gamma_{4s} := \mathrm{cum}(s(n), s^*(n), s^*(n), s(n))$ denotes the kurtosis of input $s(n)$. It follows easily that

$$\begin{aligned}&\lim_{T\to\infty}\kappa\\ &\quad= F_1(p_1)F_1^*(p_2)F_1^*(p_3)F_1(p_4)h(l_1)h^*(l_2)h^*(l_3)h(l_4)\\ &\qquad\times e^{(j2\pi(-k_1 l_2 + k_2 l_4)/P)}.\end{aligned}$$

Henceforth, we obtain that

$$\begin{aligned}\lim_{T\to\infty}\zeta_2 =& \theta(k_1, k_2)\sum_{l_2=0}^{L} h(l_2+\tau_1)h^*(l_2)e^{(-j2\pi k_1 l_2/P)}\\ &\times\sum_{l_4=0}^{L} h^*(l_4+\tau_2)h(l_4)e^{(j2\pi k_2 l_4/P)}.\end{aligned} \tag{107}$$

Applying Parseval's formula, it can be easily shown that $\lim_{T\to\infty}\zeta_2 = \xi_2$ [6]. It is also transparent that the asymptotic covariance matrix $\mathbf{\Sigma}$ is not dependent on the additional modulation factor $\exp(j\theta_e n)$ superimposed on the input $s(n)$. This is due to the fact that all the entries of the asymptotic covariance matrix contain only terms of the form $\mathbb{E}u(n)u^*(n)$. □

REFERENCES

[1] K. Abed-Meraim, J. F. Cardoso, A. Gorokhov, P. Loubaton, and E. Moulines, "On subspace methods for blind identification of SIMO FIR systems," *IEEE Trans. Signal Processing*, vol. 45, pp. 42–56, Jan. 1997.

[2] B. G. Agee, "The least-squares CMA: A new technique for rapid correction of constant modulus signal," in *Proc. ICASSP*, Tokyo, Japan, 1986, pp. 953–956.

[3] A. Belouchrani and W. Ren, "Blind carrier phase tracking with guaranteed global convergence," *IEEE Trans. Signal Processing*, vol. 45, pp. 1889–1894, July 1997.

[4] D. R. Brillinger, *Time Series Data Analysis and Theory*. San Francisco, CA: Holden Day, 1981.

[5] ——, "The comparison of least-squares and third-order periodogram procedures in the estimation of bifrequency," *J. Time Series Anal.*, vol. 1, no. 2, pp. 95–102, 1980.

[6] A. Chevreuil, "Blind equalization in a cyclostationary context," Ph.D. dissertation, ENST-Telecom Paris, Dept. Signal Processing, Paris, France, 1997.

[7] A. Chevreuil and P. Loubaton, "Blind second-order identification of FIR channels: Forced cyclo-stationarity and structured subspace method," *IEEE Signal Processing Lett.*, vol. 4, pp. 204–206, July 1997.

[8] A. Chevreuil, E. Serpedin, P. Loubaton, and G. B. Giannakis, "Performance analysis of blind channel estimators based on nonredundant periodic modulation precoders," in *Proc. ICASSP*, vol. VI, Seattle, WA, pp. 3397–3401.

[9] A. Chevreuil and P. Loubaton, "MIMO blind second-order equalization method and conjugate cyclostationarity," *IEEE Trans. Signal Processing*, vol. 47, pp. 572–578, Febr. 1999.

[10] A. Chevreuil, E. Serpedin, P. Loubaton, and G. B. Giannakis, "Blind channel identification and equalization using nonredundant periodic modulation precoders: Performance analysis," *IEEE Trans. Signal Processing*, vol. 48, pp. 1570–1586, June 2000.

[11] P. Ciblat, P. Loubaton, E. Serpedin, and G. B. Giannakis, "Performance of nondata aided carrier offset estimation for noncircular transmissions through frequency–selective channels," in *Proc. ICASSP*, Istanbul, Turkey, June 2000.

[12] A. V. Dandawaté and G. B. Giannakis, "Asymptotic theory of mixed time averages and kth order cyclic-moment and cumulant statistics," *IEEE Trans. Inform. Theory*, vol. 41, pp. 216–239, Jan. 1995.

[13] C. Georghiades and M. Moeneclaey, "Sequence estimation and synchronization from nonsynchronized samples," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1649–1657, Nov. 1991.

[14] G. B. Giannakis, "Filterbanks for blind channel identification and equalization," *IEEE Signal Processing Lett.*, vol. 4, pp. 184–187, June 1997.

[15] G. B. Giannakis and C. Tepedelenlioğlu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Proc. IEEE*, vol. 86, pp. 1969–1986, Oct. 1998.

[16] G. B. Giannakis and G. Zhou, "Harmonics in multiplicative and additive noise: Parameter estimation using cyclic statistics," *IEEE Trans. Signal Processing*, vol. 43, pp. 2217–21, Sept. 1995.

[17] F. Gini, R. Reggiannini, and U. Mengali, "The modified Cramer–Rao bound in vector parameter estimation," *IEEE Trans. Commun.*, vol. 46, pp. 52–60, Jan. 1998.

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

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

[20] E. J. Hannan, "The estimation of frequency," *J. Appl. Prob.*, vol. 10, pp. 510–519, 1973.

[21] T. Hasan, "Nonlinear time series regression for a class of amplitude modulated cosinusoids," *J. Time Series Anal.*, vol. 3, no. 2, pp. 109–122, 1982.

[22] N. Jablon, "Joint blind equalization, carrier recovery, and timing recovery for high-order QAM signal constellations," *IEEE Trans. Signal Processing*, vol. 40, pp. 1383–1397, June 1992.

[23] E. Moulines, P. Duhamel, J. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 516–25, Febr. 1995.

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

[25] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.

[26] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers—Part I: Unification and optimal designs—Part II: Blind channel estimation, synchronization and direct equalization," *IEEE Trans. Signal Processing*, vol. 47, pp. 1988–2022, July 1999.

[27] A. Scaglione, S. Barbarossa, and G. B. Giannakis, "Self-recovering equalization of time-selective fading channels using redundant filterbank precoders," in *Proc. IEEE Digital Signal Process. Workshop*, UT, Aug. 9–12, 1998.

[28] E. Serpedin and G. B. Giannakis, "Blind channel identification and equalization with modulation induced cyclostationarity," *IEEE Trans. Signal Processing*, vol. 46, pp. 1930–44, July 1998.

[29] E. Serpedin, A. Chevreuil, G. B. Giannakis, and P. Loubaton, "Non-data aided joint estimation of carrier frequency offset and channel using periodic modulation precoders: Performance analysis," in *Proc. ICASSP*, vol. V, Phoenix, AZ, Mar. 1999, pp. 2635–2638.

[30] A. Swami and M. Ghogho, "Performance analysis of cyclic estimators for harmonics in multiplicative and additive noise," in *Proc. ICASSP*, vol. IV, Seattle, WA, Mar. 1998, pp. 2309–2313.

[31] P. Stoica and R. Moses, *Introduction to Spectral Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1997.

[32] L. Tong, "Joint blind signal detection and carrier recovery over fading channel," in *Proc. ICASSP*, vol. V, Detroit, MI, May 1995, pp. 1205–1208.

[33] L. Tong, G. Xu, B. Hassibi, and T. Kailath, "Blind channel identification based on second-order statistics: A frequency-domain approach," *IEEE Trans. Inform. Theory*, vol. 41, pp. 329–334, Jan. 1995.

[34] M. K. Tsatsanis and G. B. Giannakis, "Transmitter induced cyclostationarity for blind channel equalization," *IEEE Trans. Signal Processing*, vol. 45, pp. 1785–1794, July 1997.

[35] A.-J. van der Veen, "Blind separation of BPSK sources with residual carriers," *Signal Process.*, vol. 73, pp. 67–79, Jan. 1999.

[36] G. Zhou and G. B. Giannakis, "Harmonics in multiplicative and additive noise: Performance analysis of cyclic estimators," *IEEE Trans. Signal Processing*, vol. 43, pp. 1445–60, June 1995.

Erchin Serpedin received the Diploma of electrical engineering (with highest distinction) from the Polytechnic Institute of Bucharest, Bucharest, Romania, in 1991. He received the specialization degree in signal processing and transmission of information from Supélec, Paris, France, in 1992, the M.Sc. degree from Georgia Institute of Technology, Atlanta, in 1992, and the Ph.D. degree from the University of Virginia, Charlottesville, in December 1998.

From 1993 to 1995, he was an Instructor with the Polytechnic Institute of Bucharest. In 1999, he joined the Wireless Communications Laboratory, Texas A&M University, College Station, as an Assistant Professor. His research interests lie in the areas of statistical signal processing and wireless communications.

Antoine Chevreuil was born in Caen, France, in 1971. He received the M.Sc. and Ph.D. degrees from Ecole Nationale Supérieure des Télécommunications, Paris, France, in 1994 and 1997, respectively.

In 1998, he was with the Telecommunications Department, the University of Louvain-la-Neuve, Louvain-la-Neuve, Belgium. He is currently an Assistant Professor with the Communication Systems Laboratory, University of Marne-la-Vallée, Marne-la-Vallée, France. His research interests are statistical signal processing for communications, system identification, and equalization.

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

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

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

Philippe Loubaton (M'88) was born in 1958 in Villers Semeuse, France. He received the M.Sc. and Ph.D. degrees from Ecole Nationale Supérieure des Télécommunications, Paris, France, in 1981 and 1988, respectively.

From 1982 to 1986, he was a Member of Technical Staff with Thomson-CSF/RGS, where he worked in digital communications. From 1986 to 1988, he was with the Institut National des Télécommunications as an Assistant Professor of electrical engineering. In 1988, he joined the Signal Processing Department, Ecole Nationale Supérieure des Télécommunications. Since 1995, he has been Professor of electrical engineering with Marne la Vallée University, Champs sur Marne, France. His present research interests are in statistical signal processing, and digital communications with a special emphasis on blind equalization, multi-user communication systems and multi carrier modulations.

Dr. Loubaton is an associate editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and is a member of the IEEE Signal Processing for Communications technical committee.