Performance Analysis of Blind Carrier Frequency Offset Estimators for Noncircular Transmissions Through Frequency-Selective Channels

Philippe Ciblat, Philippe Loubaton, Member, IEEE, Erchin Serpedin, and Georgios B. Giannakis, Fellow, IEEE

Abstract—This paper deals with the problem of blind estimation of the carrier frequency offset of a linearly modulated noncircular transmission through an unknown frequency-selective channel. A frequency estimator is developed based on the unique conjugate cyclic frequency of the received signal, which is equal to twice the frequency offset. Consistency and asymptotic normality of the frequency estimator together with a closed-form expression for its asymptotic variance are also established. The closed-form expression of the asymptotic variance enables analysis of the performance of the proposed frequency offset estimator as a function of the number of estimated cyclic correlation coefficients used. It is shown that optimum is obtained if the number of correlation coefficients taken into account coincides with the degree of the channel. Numerical simulations are provided and confirm the conclusion of the theoretical asymptotic analysis.

Index Terms—Asymptotic analysis, cyclic-correlation, cyclic frequency, harmonic retrieval, multiplicative noise.

I. INTRODUCTION

B LIND estimation of the carrier frequency offset and/or Doppler shifts is well motivated when it comes to compensate the local oscillator drifts and Doppler shifts induced by the relative motion of mobiles in wireless communication systems. Traditionally, channel estimation and synchronization rely on the use of a set of known symbols (sync word), whose temporal position is acquired at the receiver by cross correlating the received signal with a prestored sync word. However, this acquisition is very difficult to perform, if not impossible, for frequency-selective channels affected by frequency offset [24]. Data-aided techniques are not particularly useful for compensating the unknown intersymbol interference (ISI) effects in the presence of residual frequency-offset [24]. Therefore, developing fast-converging blind or nondata-aided carrier frequency

Manuscript received December 7, 2000; revised September 20, 2001. This work was supported by a DGA/CNRS fellowship. The associate editor coordinating the review of this paper and approving it for publication was Dr. Inbar Fijalkow.

P. Ciblat is with the Département Communications et Electronique, Ecole Nationale Supérieure des Télécommunications, Paris, France (e-mail: ciblat@com.enst.fr).

P. Loubaton is with the Laboratoire Système de Communication, Université de Marne-la-Vallée, Champs sur Marne, France (e-mail: loubaton@univmlv.fr).

E. Serpedin is with the Department of Electrical Engineering, Texas A&M University, College Station, TX 77845 USA (e-mail: erchin@spcom.tamu.edu).

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(02)00406-3.

offset estimators for channels affected by unknown ISI effects appears as an important problem.

In this paper, it is assumed that a linearly-modulated signal is transmitted through an unknown frequency-selective channel. The continuous-time received waveform $y_a(t)$ is supposed to be affected by a carrier frequency offset and/or Doppler shift F_e and is given by the following equation:

$$y_a(t) = \left(\sum_{k=-\infty}^{\infty} s_k h_a \left(t - kT_s\right)\right) e^{2i\pi F_e t} + w_a(t) \quad (1)$$

where F_e represents the carrier frequency offset, and $\{s_k\}$ denotes the independently and identically distributed (i.i.d.) symbol sequence, which is assumed of zero mean, unit variance, and *noncircularly* distributed (i.e., $\mathbb{E}[s_k^2] \neq 0$). The additive noise $w_a(t)$ is assumed normally distributed, the baud rate of the transmitter is denoted by $1/T_s$, and $h_a(t)$ stands for the convolution of the transmit and receive filters with a generally unknown multipath channel. Without any restriction, the channel $h_a(t)$ is assumed causal and time limited.

In order to retrieve the symbols s_k from a sampled version of the observation, it is necessary to estimate and compensate the multiplicative noise effect introduced by the carrier frequency offset and the additive ISI effects due to the frequency-selective channel. Channel and residual carrier frequency offset estimation is usually performed by transmitting periodically a known training sequence. However, such an approach reduces the effective transmission rate and is not feasible in many applications such as multipoint or distributed communication networks and military interception systems. It is therefore useful to explore blind solutions for estimating and compensating the carrier frequency offset. Thus far, only a few works have addressed the joint blind estimation/equalization of the channel in the presence of residual carrier (see e.g., [1]–[4] and [23]). Most of these approaches rely on a two-steps procedure: First, the channel is equalized using a constant modulus algorithm (CMA), and then, the residual carrier is tracked at the output of the equalizer. However, it is well known that this approach is successful only if the symbol sequence is circular [7]. Minimization of a kurtosis-based criterion enables channel equalization with a noncircular input constellation but only in the absence of residual carrier [7]. Therefore, the above approach is not suitable for noncircular transmissions.

In the noncircular symbol case, it is thus necessary to estimate and compensate the carrier frequency offset before equalizing the unknown channel. For this, it is possible to use the wellknown fact that $2F_e$ is the unique conjugate cyclic frequency of the received signal (see, e.g., [8]–[11]). The purpose of this paper is to establish and analyze the statistical performance of the corresponding estimation schemes and to evaluate the loss in performance relative to the scenario when the frequency offset is estimated from the equalized output of the channel, assuming perfect channel knowledge and ideal ISI cancelation.

II. CARRIER FREQUENCY OFFSET ESTIMATOR AND RELATED WORKS

Denote by y(n) the discrete-time signal obtained by sampling the continuous-time waveform $y_a(t)$ at the symbol rate $1/T_s$: $y(n) := y_a(nT_s)$. From (1), the following discrete-time channel model is obtained:

$$y(n) = \left(\sum_{l=0}^{L} h(l)s_{n-l}\right) e^{2i\pi f_e n} + w(n)$$

= ([h(z)] \cdot s_n) e^{2i\pi f_e n} + w(n) (2)

where $h(z) := \sum_{l=0}^{L} h_a(lT_s) z^{-k}$ denotes the sampled version of $h_a(t)$ at the baud rate $1/T_s$, L is the degree of the polynomial h(z), and $w(n) := w_a(nT_s)$. The discrete-time equivalent frequency offset is defined as¹ $f_e := F_e T_s \mod 1$. The estimation of f_e is thus equivalent to that of estimating F_e .

Equation (2) can be rewritten as follows:

$$y(n) = a(n)e^{2i\pi f_e n} + w(n)$$

with

$$a(n) = [h(z)] \cdot s_n.$$

Therefore, y(n) can be interpreted as a complex sinusoid corrupted by the additive noise w(n) and the *colored* multiplicative noise a(n). Frequency estimation of harmonics corrupted by additive and multiplicative noise was already studied by several authors (see, e.g., [8]–[11]). These estimators exploit the common feature that $\alpha_0 = 2f_e$ is the unique conjugate cyclic frequency of the discrete-time signal y(n). Indeed, the conjugate autocorrelation function $r_{y(c)}(n,\tau) := \mathbb{E}[y(n+\tau)y(n)]$ of y(n) can be expressed as

$$r_{y^{(c)}}(n,\tau) = r_{y^{(c)}}^{(\alpha_0)}(\tau)e^{2i\pi\alpha_0 n}$$

where $r_{y(c)}^{(\alpha)}(\tau)$ stands for the conjugate cyclic correlation of y(n) at lag τ and conjugate cyclic frequency α and is obtained as the generalized Fourier series (FS) coefficient of the time-varying correlation $r_{y(c)}(n,\tau)$:

$$\begin{aligned} r_{y^{(c)}}^{(\alpha)}(\tau) &\coloneqq \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}[y(n+\tau)y(n)] e^{-2i\pi\alpha n} \\ &= \delta\left(\alpha - \alpha_0\right) \sum_{l=-\infty}^{\infty} h_{l+\tau} h_l = \delta\left(\alpha - \alpha_0\right) r_{y^{(c)}}^{(\alpha_0)}(\tau). \end{aligned}$$

¹The expression *b* mod *a* stands for the value of *b* modulo *a*. By convention, it belongs to the following interval [-a/2, a/2].

Since for each τ , $r_{y^{(c)}}^{(\alpha)}(\tau) = 0$ when $\alpha \neq \alpha_0$, the conjugate cyclic correlation coefficients enable to retrieve $\alpha_0(f_e)$ as follows:

$$\alpha_0 = \arg \max_{\alpha \in (-0.5, 0.5)} J(\alpha), \quad J(\alpha) = \left\| \mathbf{r}_{y^{(\alpha)}}^{(\alpha)} \right\|^2$$

with $r_{y^{(c)}}^{(\alpha)} := [r_{y^{(c)}}^{(\alpha)}(-M), \dots, r_{y^{(c)}}^{(\alpha)}(M)]^{\mathrm{T}}$ and 2M + 1 denoting the number of conjugate cyclic correlation lags considered. In practice, the unknown set of correlations $\mathbf{r}_{y^{(c)}}^{(\alpha)}$ is estimated using the following consistent estimate (see, e.g., [20]):

$$\hat{\mathbf{r}}_{y^{(c)},N}^{(\alpha)} = \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{y}_2(n) e^{-2i\pi\alpha n}$$

where

$$\mathbf{y}_2(n) := [y(n-M)y(n), \dots, y(n+M)y(n)]^{\mathrm{T}}$$
. (3)

Hence, α_0 can be estimated using the estimator

$$\hat{\alpha}_N = \arg \max_{\alpha \in (-0.5, 0.5)} J_N(\alpha), \quad J_N(\alpha) = \left\| \hat{\mathbf{r}}_{y^{(c)}, N}^{(\alpha)} \right\|^2.$$
(4)

The statistical performance of this estimator was studied extensively in [8]–[11] but only in the case M = 0. As it will be shown later, the choice M = 0 is quite relevant if the multiplicative noise a(n) is white. If a(n) is colored, as is the case in the present context, it is reasonable to expect that choosing M > 0 leads to better performance.

In this paper, we prove the consistency and asymptotic normality of estimate $\hat{\alpha}_N$ for M > 0. We show that the rate of convergence of $\hat{\alpha}_N$ is $O(1/N^{3/2})$ and provides a closed-form expression for its asymptotic variance defined by

$$\gamma = \lim_{N \to \infty} N^3 \mathbb{E}\left[\left(\hat{\alpha}_N - \alpha_0 \right)^2 \right].$$
 (5)

We rely on this expression to discuss the choice of M. In particular, we show that if M is greater than the channel memory L, then the asymptotic variance is proportional to the variance of the additive noise and, thus, converges to 0 when the signal-to-noise ratio (SNR) increases. By choosing M > L, it is shown that in presence of an unknown frequency-selective channel of arbitrary memory, the asymptotic variance of the frequency offset estimator achieves almost the same asymptotic variance as the frequency offset estimator in the presence of a flat-fading channel, i.e., an ideally pre-equalized channel with no ISI effects.

The starting point of the technical part of our work is the observation that the multivariate signal $\mathbf{y}_2(n)$ can be interpreted as a (multivariate) complex sinusoid of frequency α_0 corrupted by a nonstationary additive noise and that the cost function $J_N(\alpha)$ is equivalent to a periodogram [8], [10], [12]. The standard approach to perform the asymptotic analysis of the periodogram estimates is to introduce an auxiliary nonlinear least-squares problem [13]–[16]. However, calculating the variance of $\hat{\alpha}_N$ by this approach necessitates complicated and tedious manipulations that do not lead to interpretable and closed-form expressions when M > 0. We show that the auxiliary nonlinear

²The superscript ^T denotes transposition.

least-squares criterion is not necessary. We establish the asymptotic properties of $\hat{\alpha}_N$ by using an alternative approach and obtain a closed-form expression for the asymptotic variance of $\hat{\alpha}_N$.

This paper is organized as follows. Section III establishes the link between the carrier frequency offset estimation problem and the problem of estimating the frequency of a constant amplitude harmonic embedded in additive noise. The asymptotic behavior of $\hat{\alpha}_N$ and the closed-form expression for its asymptotic variance are established in Section IV. In Section V, a theoretical analysis of the influence of the parameter M and the filter h(z) on the asymptotic performance of the frequency offset estimator is conducted. In Section VI, practical issues regarding the calculation of $\hat{\alpha}_N$ are addressed, and numerical simulations are performed in order to study the relevance of the conclusion provided by the asymptotic analysis. Finally, a conclusion is drawn in Section VII.

III. HARMONIC RETRIEVAL LINKS

In order to show the equivalence between the present carrier frequency offset estimation problem and the problem of estimating a constant amplitude harmonic embedded in noise, we first remark that

$$\mathbb{E}\left[\mathbf{y}_{2}(n)\right] = \mathbf{r}_{y^{(c)}}^{(\alpha_{0})} e^{2i\pi\alpha_{0}n}.$$
(6)

Consider the zero-mean (2M + 1)-dimensional vector $\mathbf{e}(n)$ defined by

$$\mathbf{e}(n) := \mathbf{y}_2(n) - \mathbb{E}\left[\mathbf{y}_2(n)\right]. \tag{7}$$

From (6) and (7), it follows that

$$\mathbf{y}_2(n) = \mathbf{r}_{y^{(c)}}^{(\alpha_0)} e^{2i\pi\alpha_0 n} + \mathbf{e}(n).$$
(8)

Therefore, $\mathbf{y}_2(n)$ can be interpreted as a multidimensional harmonic of frequency α_0 corrupted by the additive noise $\mathbf{e}(n)$. Moreover, the criterion $J_N(\alpha)$ in (4) is a periodogram because

$$J_N(\alpha) = \left\| \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{y}_2(n) e^{-2i\pi\alpha n} \right\|^2.$$
(9)

It is easy to check that the autocorrelation function $\mathbf{R}_{\mathbf{e}}(n, \tau) := \mathbb{E}[\mathbf{e}(n+\tau)\mathbf{e}^*(n)]$ depends only on the lag τ ($\mathbf{R}_{\mathbf{e}}(n,\tau) = \mathbf{R}_{\mathbf{e}}(\tau)$) so that $\mathbf{e}(n)$ is stationary with respect to its autocorrelation function. On the contrary, its conjugate autocorrelation function $\mathbf{R}_{\mathbf{e}^{(c)}}(n,\tau) := \mathbb{E}[\mathbf{e}(n+\tau)\mathbf{e}(n)]$ depends on τ as well as n:

$$\mathbf{R}_{\mathbf{e}^{(c)}}(n,\tau) = \mathbf{R}_{\mathbf{e}^{(c)}}^{(2\alpha_0)}(\tau)e^{2i\pi 2\alpha_0 n}.$$

Thus, $\mathbf{e}(n)$ is not a stationary process.

Thus far, numerous works have been devoted to the problem of retrieving the parameters of a number of harmonics embedded in noise by means of a periodogram estimator [13]–[16]. However, in the present paper, we deal with a more general problem. The main differences between the present context and the previously mentioned works are twofold.

- i) $\mathbf{e}(n)$ is not stationary but cyclostationary.
- ii) $\mathbf{y}_2(n)$ is a multivariate process.

However, most of the results in [13], [14], and [16] can be generalized to the present context. It is thus possible to adapt the approach of [13], [14], and [16] based on the introduction of the following nonlinear least-squares estimation (NLSE) problem:

$$\left\lfloor \hat{\theta}_N, \hat{\alpha}_N^{(K)} \right\rfloor := \arg \min_{\alpha \in (-0.5, 0.5), \theta \in \mathbb{C}^{2M+1}} K_N(\theta, \alpha)$$

where $K_N(\theta, \alpha)$ is the cost function defined by

$$K_N(\theta, \alpha) = \frac{1}{N} \sum_{n=0}^{N-1} \left\| \mathbf{y}_2(n) - \theta e^{2i\pi\alpha n} \right\|^2$$

(see, e.g., [9], [11], and [12] for the case M = 0). Consistency and asymptotic normality of the NLS-estimate $\hat{\alpha}_N^{(K)}$ are rather easy to obtain. Moreover, it can be shown that the estimates $\hat{\alpha}_N^{(K)}$ and $\hat{\alpha}_N$ are asymptotically equivalent, i.e., both have the same asymptotic variance. The evaluation of the asymptotic variance of $\hat{\alpha}_N$ is thus equivalent to that of $\hat{\alpha}_N^{(K)}$. However, calculating the asymptotic variance of $\hat{\alpha}_N^{(K)}$ is quite difficult because it requires the asymptotic covariance matrix of the vector-valued estimate $[\hat{\theta}_N, \hat{\alpha}_N^{(K)}]^{\mathrm{T}}$. Using this approach, it is quite difficult, not to say impossible, to obtain an interpretable closed-form expression for the asymptotic variance of $\hat{\alpha}_N$ when M > 0. More precisely, if M > L, it is difficult to show that the variance of $\hat{\alpha}_N$ converges to zero when the additive noise variance converges to zero.

In the next section, we will develop a quite different approach by generalizing the results shortly sketched in [15] when M = 0, and e(n) is a stationary process.

IV. ASYMPTOTIC ANALYSIS

In the sequel, an overbar will be used to denote complex conjugation. If $\mathbf{x}_1, \ldots, \mathbf{x}_L$ are random vectors, the notation cum_L ($\mathbf{x}_1, \ldots, \mathbf{x}_L$) will stand for the *L*th-order cumulant tensor of the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_L$. As the filter h(z) in (2) is FIR, $\mathbf{e}(n)$ satisfies the following mixing condition assuming very mild standard mixing assumptions on w(n) [22, p. 8, pp. 25–27].

Condition 1: If $e^{(0)}(n) = e(n)$ and $e^{(1)}(n) = \overline{e(n)}$, then we have the equation at the bottom of the page.

Essentially, the mixing Condition 1 refers to the fact that sufficiently separated samples are approximatively independent and is satisfied by all the finite memory signals encountered in practice. Condition 1 appears useful in establishing the asymptotic behavior of the proposed frequency offset estimator. Using Condition 1, it is possible to prove the following key lemma.

$$\forall L, \exists \mathcal{M}_L < \infty, \ \forall n_1, \forall (\nu_1, \dots, \nu_L) \in \{0, 1\}^L \sum_{(n_2, \dots, n_L) \in \mathbb{Z}^{L-1}} \left\| \operatorname{cum}_L \left(\mathbf{e}^{(\nu_1)}(n_1), \dots, \mathbf{e}^{(\nu_L)}(n_L) \right) \right\| \leq \mathcal{M}_L.$$

Lemma 1: Let

$$\mathbf{s}_{N}^{(K)}(\alpha) := \frac{1}{N^{K+1}} \sum_{n=0}^{N-1} n^{K} \mathbf{e}(n) e^{2i\pi\alpha n}.$$

Then3

$$\forall K \in \mathbb{N}, \ \sup_{\alpha \in [0,1]} \left\| \mathbf{s}_N^{(K)}(\alpha) \right\| \xrightarrow{a.s.} 0, \text{ as } N \to \infty.$$

This result extends the main lemma introduced in [15] to any multivariate nonstationary process that satisfies Condition 1. The proof of [15] can be generalized to our context. For more details on the technical details, see [17] and [18]. This result is important because it shows, in some sense, that the contribution of the additive noise $\mathbf{e}(n)$ in the periodogram $J_N(\alpha)$ (9) of $\mathbf{y}_2(n)$ is vanishing asymptotically while the harmonic component $\mathbf{r}_{u^{(\alpha)}}^{(\alpha_0)}e^{2i\pi\alpha_0 n}$ produces a nonzero contribution.

Using Lemma 1, one can show the following theorem.

Theorem 1: Under the mixing Condition 1, $\hat{\alpha}_N$ satisfies $\hat{\alpha}_N - \alpha_0 \xrightarrow{a.s.} 0$ and $N(\hat{\alpha}_N - \alpha_0) \xrightarrow{a.s.} 0$ as $N \to \infty$.

Proof: See Appendix A.

We now establish that $N^{3/2}\hat{\alpha}_N$ is asymptotically Gaussian. For this, we note that

$$\left. \frac{dJ_N(\alpha)}{d\alpha} \right|_{\alpha = \hat{\alpha}_N} = 0.$$

Using a first-order Taylor expansion of the derivative of $J_N(\alpha)$ around α_0 , we obtain that

$$N^{3/2}(\hat{\alpha}_N - \alpha_0) = -\mathcal{A}_N^{-1}\mathcal{B}_N \tag{10}$$

where

$$\mathcal{A}_{N} = \frac{1}{N^{2}} \left. \frac{d^{2} J_{N}(\alpha)}{d\alpha^{2}} \right|_{\alpha = \tilde{\alpha}_{N}} \tag{11}$$

$$\mathcal{B}_N = \frac{1}{\sqrt{N}} \left. \frac{dJ_N(\alpha)}{d\alpha} \right|_{\alpha = \alpha_0} \tag{12}$$

and where $\tilde{\alpha}_N$ belongs to $[\alpha_0, \hat{\alpha}_N]$ or $[\hat{\alpha}_N, \alpha_0]$. In order to analyze the asymptotic properties of $N^{3/2}(\hat{\alpha}_N - \alpha_0)$, we have to study the asymptotic behavior of \mathcal{A}_N and \mathcal{B}_N . The following two theorems are proved in Appendices B and C, respectively.

Theorem 2: Assuming the mixing Condition 1, the following relationship holds:

$$\mathcal{A}_N \xrightarrow{a.s.} \gamma_{\mathcal{A}} = -\frac{2\pi^2}{3} \left\| \mathbf{r}_{y^{(c)}}^{(\alpha_0)} \right\|^2 \quad as \quad N \to \infty.$$

Theorem 3: Assuming the mixing Condition 1, it follows that $\mathcal{B}_N \xrightarrow{\mathcal{L}} \mathcal{N}(0, \gamma_{\mathcal{B}})$ in distribution as $N \to \infty$, where the expression of $\gamma_{\mathcal{B}}$ can be deduced from the calculations presented in Appendix C.

Using (10) and Theorems 2 and 3, we obtain the main result of this section.

³Notation *a.s.* stands for *almost surely convergence* or *convergence with probability one.*

Theorem 4: Assuming the mixing Condition 1, it follows that $N^{3/2}(\hat{\alpha}_N - \alpha_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \gamma)$ in distribution as $N \to \infty$ where $\gamma := \gamma_{\mathcal{A}}^{-1} \gamma_{\mathcal{B}} \gamma_{\mathcal{A}}^{-1}$ is given by

$$\gamma = \frac{3}{\pi^2} \frac{\mathbf{R}_{y^{(c)}}^{(\alpha_0)^*} \mathbf{G} \mathbf{R}_{y^{(c)}}^{(\alpha_0)}}{\left\| \mathbf{R}_{y^{(c)}}^{(\alpha_0)} \right\|^4}$$
(13)

and

$$\mathbf{R}_{y^{(c)}}^{(\alpha_0)} = \begin{bmatrix} \mathbf{r}_{y^{(c)}}^{(\alpha_0)} \\ \mathbf{r}_{y^{(c)}}^{(\alpha_0)} \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} \mathbf{\Gamma} & -\mathbf{\Gamma} \\ -\mathbf{\Gamma}^{(c)} & \mathbf{\overline{\Gamma}} \end{bmatrix}$$

Matrices Γ and $\Gamma^{(c)}$ denote the unconjugated/conjugated asymptotic covariance matrices of $\delta \hat{\mathbf{r}}_{y^{(c)},N}^{(\alpha_0)} := \hat{\mathbf{r}}_{y^{(c)},N}^{(\alpha_0)} - \mathbf{r}_{y^{(c)}}^{(\alpha_0)}$, i.e.,

$$\boldsymbol{\Gamma} \coloneqq \lim_{N \to \infty} N \mathbb{E} \left[\delta \hat{\mathbf{r}}_{y^{(c)},N}^{(\alpha_0)} \delta \hat{\mathbf{r}}_{y^{(c)},N}^{(\alpha_0)^*} \right]$$
$$\boldsymbol{\Gamma}^{(c)} \coloneqq \lim_{N \to \infty} N \mathbb{E} \left[\delta \hat{\mathbf{r}}_{y^{(c)},N}^{(\alpha_0)} \delta \hat{\mathbf{r}}_{y^{(c)},N}^{(\alpha_0)\mathrm{T}} \right]$$

Proof: See Appendix E.

We have thus proved the asymptotic normality of $N^{3/2}(\hat{\alpha}_N - \alpha_0)$ and that the convergence rate of $(\hat{\alpha}_N - \alpha_0)$ is $N^{3/2}$, as encountered in standard constant amplitude harmonic retrieval problems.

V. INFLUENCE OF M AND h(z)

Although it is possible to evaluate γ in closed-form in the general noncircular case, we only focus in the following on the behavior of the estimator when $\{s_n\}_{n \in \mathbb{Z}}$ is real valued. This is because the most powerful result of this section only holds in the real-valued case. In practice, this limitation is not very restrictive because the noncircular constellations are very often real valued.

In order to analyze the influence of the parameter M on γ , we first have to derive the closed-form expressions of matrices Γ and $\Gamma^{(c)}$. For this, we first introduce the following notations. We denote by $S_y(\exp(2i\pi f))$ and $S_w(\exp(2i\pi f))$ the spectral densities of y(n) and w(n), respectively. If σ^2 represents the variance of the noise, $S_w(\exp(2i\pi f))$ can be expressed as $S_w(\exp(2i\pi f)) = \sigma^2 |g(\exp(2i\pi f))|^2$ for some function $g(\exp(2i\pi f))$ satisfying $\int_{-1/2}^{1/2} |g(\exp(2i\pi f))|^2 df = 1$. Note that

$$S_y\left(e^{2i\pi f}\right) = h\left(e^{2i\pi(f-(\alpha_0/2))}\right)h\left(e^{2i\pi(f-(\alpha_0/2))}\right)^* + S_w\left(e^{2i\pi f}\right).$$

Denote by $S_{y^{(c)}}^{(\alpha_0)}(\exp{(2i\pi f)})$ the conjugate cyclic spectrum of y(n) at cyclic frequency α_0 . Straightforward calculations show that

$$S_{y^{(\alpha)}}^{(\alpha_0)}\left(e^{2i\pi f}\right) = h\left(e^{2i\pi(f-\alpha_0/2)}\right)h\left(e^{-2i\pi(f-\alpha_0/2)}\right) .$$
(14)

Relying on the (28) and (29), obtained in Appendix E and using standard calculations, one can immediately obtain the following lemma.

Lemma 2: Assume that $\{s_n\}_{n \in \mathbb{Z}}$ is real valued. Then, matrices Γ and $\Gamma^{(c)}$ are given by

$$\begin{split} \mathbf{\Gamma} &= \int_{0}^{1} S_{y} \left(e^{2i\pi f} \right) S_{y} \left(e^{2i\pi(\alpha_{0} - f)} \right) \\ &\cdot \mathbf{d}_{M} \left(e^{2i\pi f} \right) \mathbf{d}_{M} \left(e^{2i\pi f} \right)^{*} df \\ &+ \int_{0}^{1} S_{y} \left(e^{2i\pi f} \right) S_{y} \left(e^{2i\pi(\alpha_{0} - f)} \right) \\ &\cdot \mathbf{d}_{M} \left(e^{2i\pi f} \right) \mathbf{d}_{M} \left(e^{2i\pi(\alpha_{0} - f)} \right)^{*} df \\ &+ \kappa \int_{0}^{1} h \left(e^{2i\pi f} \right) h \left(e^{-2i\pi f} \right) \mathbf{d}_{M} \left(e^{2i\pi(f + \alpha_{0}/2)} \right) df \\ &\cdot \int_{0}^{1} h \left(e^{2i\pi f} \right)^{*} h \left(e^{-2i\pi f} \right)^{*} \mathbf{d}_{M} \left(e^{2i\pi(f + \alpha_{0}/2)} \right)^{*} df \\ \mathbf{\Gamma}^{(c)} &= \int_{0}^{1} S_{y^{(c)}}^{(\alpha_{0})} \left(e^{2i\pi f} \right) S_{y^{(c)}}^{(\alpha_{0})} \left(e^{2i\pi(\alpha_{0} - f)} \right) \\ &\cdot \mathbf{d}_{M} \left(e^{2i\pi f} \right) \mathbf{d}_{M} \left(e^{2i\pi f} \right)^{T} df \\ &+ \int_{0}^{1} S_{y^{(c)}}^{(\alpha_{0})} \left(e^{2i\pi f} \right) S_{y^{(c)}}^{(\alpha_{0})} \left(e^{2i\pi(\alpha_{0} - f)} \right) \\ &\cdot \mathbf{d}_{M} \left(e^{2i\pi f} \right) \mathbf{d}_{M} \left(e^{2i\pi f} \right)^{T} df \\ &+ \kappa \int_{0}^{1} h \left(e^{2i\pi f} \right) h \left(e^{-2i\pi f} \right) \mathbf{d}_{M} \left(e^{2i\pi(f + \alpha_{0}/2)} \right)^{T} df \\ &\cdot \int_{0}^{1} h \left(e^{2i\pi f} \right) h \left(e^{-2i\pi f} \right) \mathbf{d}_{M} \left(e^{2i\pi(f + \alpha_{0}/2)} \right)^{T} df \end{split}$$

where $\kappa := \operatorname{cum}_4(s_n, s_n, s_n, s_n)$ is the kurtosis of s_n , and $\mathbf{d}_M(e^{2i\pi f})$ represents the (2M + 1)-dimensional vector $\mathbf{d}_M(e^{2i\pi f}) := [e^{-2i\pi M f}, \dots, e^{2i\pi M f}]^{\mathrm{T}}$.

After some straightforward but tedious manipulations, it is possible to show the following result.

Theorem 5: Assume that $\{s_n\}_{n \in \mathbb{Z}}$ is real valued; then, the asymptotic variance γ can be written as

$$\gamma = \Phi(M) + \frac{3\sigma^2}{\pi^2} \frac{Q_1 + \sigma^2 Q_2}{Q_3}$$
(15)

where $\Phi(M)$ does not depend on σ^2 nor on $g(e^{2i\pi f})$ and where Q_1, Q_2, Q_3 are given by

$$Q_{1} := \int_{0}^{1} \left(\left| g\left(e^{2i\pi f}\right) h\left(e^{-2i\pi(f-\alpha_{0}/2)}\right) \right|^{2} + \left| g\left(e^{-2i\pi(f-\alpha_{0})}\right) h\left(e^{2i\pi(f-\alpha_{0}/2)}\right) \right|^{2} \right) \\ \cdot \left| S_{y^{(c)}}^{(\alpha_{0})}\left(e^{2i\pi f}\right) \right|^{2} df$$

$$Q_{2} := \int_{0}^{1} \left| g\left(e^{2i\pi f}\right) g\left(e^{-2i\pi(f-\alpha_{0})}\right) \right|^{2} \left| S_{y^{(c)}}^{(\alpha_{0})}\left(e^{2i\pi f}\right) \right|^{2} df$$

$$Q_{3} := \left(\int_{0}^{1} \left| S_{y^{(c)}}^{(\alpha_{0})}\left(e^{2i\pi f}\right) \right|^{2} df \right)^{2}.$$
Finally,

Finally

$$\Phi(M) = 0 \quad \text{if} \quad M \ge L.$$

This result shows that if $M \ge L$, the asymptotic variance γ converges toward 0 as $\sigma^2 \rightarrow 0$. Moreover, in the noiseless case,

 $\gamma = 0$ if $M \ge L$, which means that $N^{3/2}(\hat{\alpha}_N - \alpha_0)$ converges toward 0 in probability. Therefore, the estimate $\hat{\alpha}_N$ converges faster than $N^{-3/2}$ toward α_0 if $M \ge L$. However, the estimate $\hat{\alpha}_N$ is not deterministic in the noiseless case unless h(z) is reduced to a constant, i.e., $\hat{\alpha}_N \ne \alpha_0$ for N finite. It can be shown, in particular, that the first-order derivative $(dJ_N(\alpha)/d\alpha|_{\alpha=\alpha_0})$ of $J_N(\alpha)$ at $\alpha = \alpha_0$ is not zero. We note, however, that if h(z)is reduced to a constant h_0 , then $J_N(\alpha)$ can be written as

$$J_N(\alpha) = \left| \frac{1}{N} \sum_{n=0}^{N-1} s_n^2 e^{2i\pi(\alpha_0 - \alpha)n} \right|^2.$$

As $s_n^2 \ge 0$, the argument of the minimum of $J_N(\alpha)$ clearly coincides with α_0 , i.e., $\hat{\alpha}_N = \alpha_0$. In other words, the estimator $\hat{\alpha}_N$ is deterministic if h(z) is reduced to a constant.

It is also interesting to conform expression (15) with the formulas given in [8]–[10] in the case T = 0, h(z) = 1 (the multiplicative noise is white) and the additive noise is white. In this particular case, (15) gives

$$\gamma = \frac{6}{\pi^2} \left(\sigma^2 + \frac{1}{2} \sigma^4 \right)$$

which coincides with the formulas presented in [8]–[10].

In the case where the noise is white (i.e., $g(e^{2i\pi f}) = 1$ for each f), the closed-form expression of γ can be simplified as follows:

$$\gamma = \Phi(M) + \frac{3\sigma^2}{\pi^2} \left(2 \frac{\int_0^1 |h(e^{2i\pi f})|^4 |h(e^{-2i\pi f})|^2 df}{\left(\int_0^1 |h(e^{2i\pi f})|^2 |h(e^{-2i\pi f})|^2 df\right)^2} + \frac{\sigma^2}{\int_0^1 |h(e^{2i\pi f})|^2 |h(e^{-2i\pi f})|^2 df} \right).$$
(16)

It is interesting to study the behavior of (16) when $\int_0^1 |h(e^{2i\pi f})|^2 |h(e^{-2i\pi f})|^2 df \rightarrow 0$, i.e., if the function $f \rightarrow h(e^{2i\pi f})h(e^{-2i\pi f})$ is nearly identically zero. In this case, the conjugate cyclic spectrum (14) is, of course, nearly zero (note, in particular, that $\int_0^1 |h(e^{2i\pi f})|^2 |h(e^{-2i\pi f})|^2 df = \int_0^1 |S_{y^{(c)}}^{(\alpha_0)}(e^{2i\pi f})|^2 df$). A careful analysis of the term $\Phi(M)$ shows that $\Phi(M)$ remains bounded if $\int_0^1 |h(e^{2i\pi f})|^2 |h(e^{-2i\pi f})|^2 df \rightarrow 0$. The dominant term in (16) is therefore the contribution of the noise, which converges to ∞ . The boundness of $\Phi(M)$ when $\int_0^1 |h(e^{2i\pi f})|^2 |h(e^{-2i\pi f})|^2 df \rightarrow 0$ seems to be a paradox. In the noiseless case, this tends to indicate that the performance of the estimate is insensitive to the power of the conjugate cyclic spectrum at frequency α_0 . However, the values of N for which the asymptotic analysis becomes relevant increases when $\int_0^1 |h(e^{2i\pi f})|^2 |h(e^{-2i\pi f})|^2 df \rightarrow 0$. We also note that if $\int_0^1 |h(e^{2i\pi f})|^2 |h(e^{-2i\pi f})|^2 df = 0$, the asymptotic analysis does not make sense.

VI. SIMULATED PERFORMANCE

We assume BPSK modulation for the transmitted input symbol stream. The channel h(z) is the convolution of the transmitter and receiver filters, which are square-root raised-cosine filters with roll-off factor $\rho = 0.2$, with an unknown multipath channel. The complex amplitudes of the paths are Gaussian distributed and the time delays are uniform distributed in $[0, 3T_s]$. The probability distribution of the time delays is chosen to satisfy the condition $L \leq 20$. We only consider the white noise case, i.e., w(n) is white noise. Unless otherwise stated, we average the results over 1000 realizations of the estimate $\hat{\alpha}_N$. Note the channel is modified at each trial.

The cost function $J_N(\alpha)$ shows, in practice, several local spurious extrema (see Fig. 1). $\hat{\alpha}_N$ is thus computed in two steps. In the first step (coarse search), the function $J_N(\alpha)$ is evaluated by means of an FFT algorithm on the grid $\{a_k = k/N, \text{ for } 0 \leq k < N\}$. In the second step (fine search), a gradient minimization algorithm of $J_N(\alpha)$, initialized at the argument of $\min_{a_k} J(a_k)$, gives the estimate $\hat{\alpha}_N$.

As we show later, the performance of the estimate is satisfying if the coarse search selects the frequency a_k of the grid, denoted $a_{k_{opt}}$, which is the closest to α_0 . As the performance of the first step is, of course, not predicted by the asymptotic analysis, we briefly study its performance.

A. Experimental Study of the First Step

We evaluate the values N versus SNR for which the percentage of detections of frequency $a_{k_{opt}}$ is equal to 0.99. Fig. 2 is evaluated by averaging on 5000 trials. It shows that the choice M = L outperforms significantly the choice M = 0. For M = L, the value N = 400 provides satisfying detection rate if SNR ≥ 10 dB.

We have observed that wrong detections occur when the power of the conjugate cyclic spectrum at frequency α_0 is low. In this case, second-order cyclic methods obviously fail.

B. Experimental Study of the Second Step

We study the behavior of the mean square error of $\hat{\alpha}_N$ with respect to the SNR, the number of observations N, and the design parameter M, respectively. Our purpose is twofold. On one hand, we wish to confirm the accuracy of the theoretical asymptotic analysis. For this, we compare the theoretical and empirical mean square errors. The empirical mean square errors are obtained by averaging, over MC = 1000 Monte Carlo trials, the estimation errors $(\hat{\alpha}_N - \alpha_0)^2 / \alpha_0^2$. On the other hand, in order to study the effect of the channel h(z) on the performance, we also evaluate the performance of $\hat{\alpha}_N$ in the ideal situation of a channel that is perfectly equalized, i.e., no ISI effects are present [h(z) = 1].

The mean-square error is, of course, deeply influenced by possible wrong detections in the coarse search step. This is illustrated in Fig. 3, in which we compare theoretical and empirical values of the mean-square error versus SNR for N = 500. The empirical and theoretical mean square errors are represented by dashed lines and solid lines, respectively. We compute the mean square errors for M = 0 and M = L. We also plot the mean-square errors in the case where h(z) is reduced to a constant.

Except in the ISI-less case, the empirical results are very far from the asymptotic predictions because, over 1000 trials, one to





Fig. 2. Threshold of right detection versus SNR (MC = 5000).



Fig. 3. MSE versus SNR (N = 500, wrong detection taken into account).

five wrong detections (depending on the SNR and on the value of M) occurred.

In the next experiments, however, we show that the empirical and theoretical results coincide if we do not take into account in the mean-square errors the trials corresponding to the wrong detections. The parameters used in Fig. 4 are the same as those in Fig. 3. However, Fig. 4 represents the above-mentioned empirical modified mean-square errors.

For low SNRs, the (modified) empirical curves are closely matched by the theoretical curves. In contrast, for high SNRs, we notice that for the case M = L, the theoretical and empirical curves mismatch. This is because the $\hat{\alpha}_N$ is not a deterministic estimate. If $\sigma^2 = 0$, $N^{3/2}(\hat{\alpha}_N - \alpha_0)$ converges in probability to 0, but for a given value of N, $\hat{\alpha}_N$ never coincides with α_0 . If σ^2 is too small, the term of $\hat{\alpha}_N - \alpha_0$ converging faster than $N^{3/2}$ may provide a residual variance greater than theoretical asymptotic variance. In the ISI-less case, $\hat{\alpha}_N$ is deterministic, and the empirical and theoretical curves match, whatever the SNR.

In Fig. 5, the theoretical and modified empirical mean square errors of the proposed frequency offset estimator are plotted versus N, assuming a fixed SNR = 20 dB.

The theoretical and empirical curves for M = 0 are very close one to the other for $N \ge 300$. The number of samples N must be chosen larger ($N \ge 500$) to observe a good fit between the theoretical and empirical curves for M = L. However, the empirical performances for M = L outperform quite significantly those of the estimate corresponding to M = 0 for any value of N.

Finally, we analyze the influence of M on the asymptotic of the carrier frequency estimator. We plot in Fig. 6 the theoretical and modified empirical mean square errors versus M for SNR = 20 dB and N = 500.

In this particular case, a significant improvement is observed as soon as $M \ge 4$.

VII. CONCLUSIONS

In this paper, we have analyzed the performance of a frequency offset estimator assuming an unknown frequency-selective channel and noncircularly distributed symbols. We have shown that the proposed estimator is consistent and asymptotically normal and that its convergence rate is $O(1/N^{3/2})$. We have expressed its asymptotic variance in closed form and have analyzed the influence of the number of cyclo-correlations (2M+1) on the performance of the frequency offset estimator. We have shown that choosing a number of correlations greater than the memory of the channel leads to a quite significant improvement. In practice, the cost function to be maximized shows spurious local extrema. Therefore, the estimate is calculated using a coarse search on a FFT grid followed by a gradient algorithm properly initialized. As the coarse search may fail to select a good initialization, the empirical results do not match the theoretical ones. However, if the trials on which wrong detections occur are not taken account in the evaluation of the mean-square errors, the theoretical and empirical results are in agreement. Finally, it is interesting to note that most of the results of this paper are valid if the signal $a(n) = [h(z)]s_n$ is replaced by a more general noncircular multiplicative noise. The most important conclusion of Theorem 5 ($\Phi(M) = 0$ if M > L), however, remains valid if the multiplicative noise is modeled as the output of an FIR filter driven by a real i.i.d sequence.



Fig. 4. MSE versus SNR (N = 500, only right detection taken into account).







Fig. 6. MSE versus M (SNR = 20 dB, N = 500).

APPENDIX A PROOF OF CONSISTENCY OF THE CYCLIC FREQUENCY ESTIMATOR $\hat{\alpha}_N$

For this, we first note that the sequence $\{\hat{\alpha}_N\}_{N\geq 0}$ belongs to the compact set $\mathcal{I} \subset (-0.5, 0.5)$. Therefore, in order to prove that $\hat{\alpha}_N$ converges almost surely to α_0 , it is sufficient to establish that every convergent subsequence extracted from $\hat{\alpha}_N$ converges to α_0 . We consider a subsequence $\{\hat{\alpha}_{\phi(N)}\}_{N\in\mathbb{N}}$ converging to a certain value $\alpha_1 \in \mathcal{I}$, and we will prove that $\alpha_1 = \alpha_0$.

Since $\hat{\alpha}_N$ maximizes $J_N(\alpha)$, it follows that $J_N(\hat{\alpha}_N) \geq J_N(\alpha_0)$. Moreover, this inequality still holds for the subsequence, and we immediately obtain that

$$\Delta J = \lim_{N \to \infty} \left[J_{\phi(N)} \left(\hat{\alpha}_{\phi(N)} \right) - J_{\phi(N)}(\alpha_0) \right]$$

exists almost surely and is non-negative (i.e., $\Delta J \ge 0$).

One observes that $J_N(\alpha)$ can be decomposed into four terms:

$$J_N(\alpha) = \mathbf{t}_N^*(\alpha)\mathbf{t}_N(\alpha) + \mathbf{s}_N^{(0)*}(\alpha)\mathbf{t}_N(\alpha) + \mathbf{t}_N^*(\alpha)\mathbf{s}_N^{(0)}(\alpha) + \mathbf{s}_N^{(0)*}(\alpha)\mathbf{s}_N^{(0)}(\alpha)$$

with $\mathbf{t}_N(\alpha) := (1/N) \sum_{n=0}^{N-1} \mathbf{r}_{y^{(c)}}(n) \exp(-2i\pi\alpha n)$ and $\mathbf{s}_N^{(0)}(\alpha)$, which is defined in Lemma 1. Furthermore, $\mathbf{t}_N(\alpha)$ is bounded with respect to N and α . According to Lemma 1, it follows that

$$\lim_{N \to \infty} J_{\phi(N)}(\alpha_0) \stackrel{a.s.}{=} \lim_{N \to \infty} \left\| \mathbf{t}_{\phi(N)}(\alpha_0) \right\|^2$$
(17)

$$\lim_{N \to \infty} J_{\phi(N)}(\hat{\alpha}_{\phi(N)})^{a.s.} \lim_{N \to \infty} \left\| \mathbf{t}_{\phi(N)}(\hat{\alpha}_{\phi(N)}) \right\|^2.$$
(18)

Since $\mathbf{r}_{y^{(c)}}(n) = \mathbf{r}_{y^{(c)}}^{(\alpha_0)} \exp{(2i\pi\alpha_0 n)}$, it follows that

$$\lim_{N \to \infty} J_{\phi(N)}(\alpha_0) \stackrel{a.s.}{=} \left\| \mathbf{r}_{y^{(c)}}^{(\alpha_0)} \right\|^2.$$

In order to evaluate the limit (18), we need to introduce the following lemma (see [16]).

Lemma 3

Let $\{c_N\}_{N\in\mathbb{N}}$ be a real-valued sequence, belonging to a compact set that is included in (-1/2, 1/2] and converging to c. Define

$$q_N(c_N) := \frac{1}{N} \sum_{n=0}^{N-1} e^{2i\pi c_N n}.$$

Then, as $N \to +\infty$, the following relations hold.

• $q_N(c_N) \rightarrow 0$ if $c \neq 0$.

•
$$q_N(c_N) \to 0$$
 if $c = 0$ and $N|c_N - c| \to \infty$

•
$$q_N(c_N) \to e^{i\beta} \operatorname{sin} c(\beta)$$
 if $c = 0$ and $N(c_N - c) \to \beta \in \mathbb{R}$

One can check that

$$\lim_{N \to \infty} J_{\phi(N)} \left(\hat{\alpha}_{\phi(N)} \right) \stackrel{a.s.}{=} \lim_{N \to \infty} \left\| \mathbf{r}_{y^{(c)}}^{(\alpha_0)} \right\|^2 \cdot \left| q_{\phi(N)} \left(\alpha_0 - \hat{\alpha}_{\phi(N)} \right) \right|^2.$$

If $\alpha_1 \neq \alpha_0$, Lemma 3 implies that $\Delta J \stackrel{a.s.}{=} - \left\| \mathbf{r}_{y^{(c)}}^{(\alpha_0)} \right\|^2 < 0$, which contradicts the condition $\Delta J \ge 0$. Therefore, $\alpha_1 = \alpha_0$.

We now consider the sequence $\{b_N\}_{N\in\mathbb{N}}$ defined by $b_N := N(\hat{\alpha}_N - \alpha_0)$. If $\{b_N\}_{N\in\mathbb{N}}$ is not bounded, there exists a subsequence $\{b_{\phi(N)}\}_{N\geq 0}$ such that $|b_{\phi(N)}| \to +\infty$. According to Lemma 3, the corresponding ΔJ is equal to $-\left\|\mathbf{r}_{y^{(c)}}^{(\alpha_0)}\right\|^2$, which

is impossible. Thus, $\{b_N\}_{N\in\mathbb{N}}$ is bounded. Let $\{b_{\phi(N)}\}_{N\in\mathbb{N}}$ be a subsequence converging to β . Using Lemma 3, we obtain that $\Delta J = \left\|\mathbf{r}_{y^{(c)}}^{(\alpha_0)}\right\|^2 (\operatorname{sin} c^2(\beta) - 1)$. As $\Delta J \ge 0, \beta$ must be equal to 0.

APPENDIX B PROOF OF THEOREM 2

We can immediately check that⁴

$$\begin{split} \mathcal{A}_{N} = & \frac{2}{N^{2}} \left\| \left. \frac{\partial \hat{\mathbf{r}}_{c,N}^{(\alpha)}}{\partial \alpha} \right|_{\alpha = \tilde{\alpha}_{N}} \right\|^{2} \\ &+ \frac{2}{N^{2}} \Re \mathbf{e} \left[\left. \frac{\partial^{2} \hat{\mathbf{r}}_{c,N}^{(\alpha)}}{(\partial \alpha)^{2}} \right|_{\alpha = \tilde{\alpha}_{N}}^{*} \cdot \left. \hat{\mathbf{r}}_{c,N}^{(\alpha)} \right|_{\alpha = \tilde{\alpha}_{N}} \right]. \end{split}$$

As $|\tilde{\alpha}_N - \alpha_0| \leq |\hat{\alpha}_N - \alpha_0|$, Theorem 1 implies that $N(\tilde{\alpha}_N - \alpha_0) \xrightarrow{a.s.} 0$. Using Lemmas 1 and 3 from Appendix A, it follows that

$$\frac{1}{N^K} \left. \frac{\partial^K \hat{\mathbf{r}}_{c,N}^{(\alpha)}}{(\partial \alpha)^K} \right|_{\alpha = \tilde{\alpha}_N} \to \frac{(-2i\pi)^K}{K+1} \mathbf{r}_{y^{(c)}}^{(\alpha_0)}$$

for any integer K. This proves Theorem 2.

APPENDIX C PROOF OF THEOREM 3

Using (8), it follows that \mathcal{B}_N can be written as

$$\mathcal{B}_N = \mathcal{B}_N^{(1)} + \mathcal{B}_N^{(2)} \tag{19}$$

where

$$\mathcal{B}_{N}^{(1)} = 2i\pi \frac{N-1}{N} \mathbf{R} \mathbf{E}_{N}$$
(20)

$$\mathcal{B}_N^{(2)} = -4\pi \Im \mathbf{m} \left[\mathbf{E}_N^*(1) \mathbf{s}_N^{(0)}(\alpha_0) \right]$$
(21)

 $\mathbf{s}_N^{(0)}(\alpha_0)$ is defined in Lemma 1, and

$$\mathbf{R} := \left[-\frac{\mathbf{r}_y^{(\alpha_0)^*}}{2}, \mathbf{r}_y^{(\alpha_0)^*}, \frac{\mathbf{r}_y^{(\alpha_0)^{\mathrm{T}}}}{2}, -\mathbf{r}_y^{(\alpha_0)^{\mathrm{T}}} \right]$$
$$\mathbf{E}_N := [\mathbf{E}_N^{\mathrm{T}}(0), \mathbf{E}_N^{\mathrm{T}}(1), \mathbf{E}_N^*(0), \mathbf{E}_N^*(1)]^{\mathrm{T}}.$$

For K = 0, 1, $\mathbf{E}_N(K)$ is defined by

$$\mathbf{E}_N(K) := \frac{1}{N^K \sqrt{N}} \sum_{n=0}^{N-1} \mathbf{e}(n) n^K e^{-2i\pi\alpha_0 n}$$

It is proved in Appendix D that \mathbf{E}_N converges in distribution to a zero-mean Gaussian distribution. This result shows that $\mathcal{B}_N^{(1)}$ is asymptotically zero-mean Gaussian. Moreover, Lemma 1 shows that $\mathbf{s}_N^{(0)}$ converges almost surely to zero. As $\mathbf{E}_N(1)$ is asymptotically Gaussian, it is bounded in probability. Therefore, $\mathcal{B}_N^{(2)} = -4\pi\Im[\mathbf{E}_N^*(1)\mathbf{s}_N^{(0)}(\alpha_0)]$ converges almost surely to zero [19]. Theorem 3 follows now immediately from (19).

 ${}^4\text{The notations}\,\Re e[.]$ and $\Im m[.]$ denote the real and imaginary parts of a complex number, respectively.

$\begin{array}{c} \text{Appendix} \ \ D \\ \text{Proof of Asymptotic Normality of } \mathbf{E}_{N} \end{array}$

Let $\operatorname{cum}_L(\mathbf{E}_N)$ denote the *L*th-order cumulant tensor of \mathbf{E}_N . We recall that $\mathbf{e}(n)$ is a (2M + 1)-multivariate process. Therefore, one can set

$$\mathbf{e}(n) = [e_{-M}(n), \dots, e_M(n)]^{\mathrm{T}}.$$

One can see that the generic form of the components of the tensor $\operatorname{cum}_L(\mathbf{E}_N)$ is given by

$$\mathcal{C}_{N}^{(L)} := N^{-\frac{L}{2}} \sum_{n_{0},\dots,n_{L-1}=0}^{N-1} D^{(\nu_{0})}(n_{0})\dots D^{(\nu_{L-1})}(n_{L-1}) \\ \cdot \operatorname{cum}_{L} \left\{ e_{\tau_{0}}^{(\nu_{0})}(n_{0}),\dots, e_{\tau_{L-1}}^{(\nu_{L-1})}(n_{L-1}) \right\}$$
(22)

where τ_l and ν_l belong to the sets $\{-T, \ldots, T\}$ and $\{0, 1\}$, respectively

$$\begin{split} D^{(0)}(n) &:= \begin{cases} e^{-2i\pi\alpha_0 n} \\ \frac{n}{N}e^{-2i\pi\alpha_0 n} \\ 0 \\ \text{and:} \\ e^{(0)}_{\tau}(n) &:= e_{\tau}(n), \quad e^{(1)}_{\tau}(n) := \overline{e_{\tau}(n)}. \end{split}$$

It follows that for $\forall \nu \in \{0,1\}$ and $\forall n \leq N$, $|D^{(\nu)}(n)| \leq 1$. Due to the triangular inequality applied on (22), we obtain (23), shown at the bottom of the page. Condition 1 implies that there is a constant \mathcal{M}_L , depending only on L, such that

$$\left|\mathcal{C}_{N}^{(L)}\right| \leq \mathcal{M}_{L} N^{-(L/2-1)}$$

Therefore, $\operatorname{cum}_L(\mathbf{E}_N) = \mathcal{O}(N^{-(L/2-1)})$. If L > 2, then (L/2-1) > 0 and it follows that

$$\lim_{N\to\infty}\operatorname{cum}_L(\mathbf{E}_N)=0$$

which implies that \mathbf{E}_N converges to a Gaussian distribution.

APPENDIX E PROOF OF THEOREM 4

As \mathcal{A}_N converges almost surely toward $\gamma_{\mathcal{A}}$ and \mathcal{B}_N converges in distribution to a centered normal distribution of variance $\gamma_{\mathcal{B}}$, $N^{3/2}(\hat{\alpha}_N - \alpha_0) = -\mathcal{A}_N^{-1}\mathcal{B}_N$ converges to a normal distribution $\mathcal{N}(0,\gamma)$ of zero-mean and standard deviation $\gamma := \gamma_{\mathcal{A}}^{-2}\gamma_{\mathcal{B}}$ (see [19]). In order to complete the proof of the theorem, we still need to establish the expression of γ . According to Theorem 3 and (20), it is easy to check that

$$\gamma_{\mathcal{B}} = 4\pi^2 \mathbf{R} \left(\lim_{N \to \infty} \mathbb{E} \left[\mathbf{E}_N \mathbf{E}_N^* \right] \right) \mathbf{R}^*.$$
 (24)

In order to evaluate $\gamma_{\mathcal{B}}$, we need to obtain a closed-form expression for $\lim_{N\to\infty} \mathbb{E}[\mathbb{E}_N \mathbb{E}_N^*]$. For this, one can observe that the matrix $\mathbb{E}[\mathbb{E}_N \mathbb{E}_N^*]$ can be expressed as

$$\begin{bmatrix} \mathbf{P}_{N}(0,0) & \mathbf{P}_{N}(0,1) & \mathbf{P}_{N}^{(c)}(0,0) & \mathbf{P}_{N}^{(c)}(0,1) \\ \\ \mathbf{P}_{N}^{(1,0)} & \mathbf{P}_{N}^{(1,1)} & \mathbf{P}_{N}^{(c)}(1,0) & \mathbf{P}_{N}^{(c)}(1,1) \\ \\ \mathbf{P}_{N}^{(c)}(0,0) & \mathbf{P}_{N}^{(c)}(0,1) & \mathbf{P}_{N}(0,0) & \mathbf{P}_{N}(0,1) \\ \\ \mathbf{P}_{N}^{(c)}(1,0) & \mathbf{P}_{N}^{(c)}(1,1) & \mathbf{P}_{N}(1,0) & \mathbf{P}_{N}(1,1) \end{bmatrix}$$
(25)

where

$$\begin{aligned} \mathbf{P}_{N}(K,K') =& \mathbb{E}\left[\mathbf{E}_{N}(K)\mathbf{E}_{N}^{*}(K')\right] \\ &= \frac{1}{N^{K+K'+1}} \\ &\cdot \sum_{\substack{n=0\\n'=0}}^{N-1} \mathbb{E}\left[\mathbf{e}(n)\mathbf{e}^{*}(n')\right] n^{K}n'^{K'}e^{2i\pi\alpha_{0}(n'-n)} \\ &\text{and:} \\ \mathbf{P}_{N}^{(c)}(K,K') =& \mathbb{E}\left[\mathbf{E}_{N}(K)\mathbf{E}_{N}^{\mathrm{T}}(K')\right] \\ &= \frac{1}{N^{K+K'+1}} \\ &\cdot \sum_{\substack{n=0\\n'=0}}^{N-1} \mathbb{E}\left[\mathbf{e}(n)\mathbf{e}^{\mathrm{T}}(n')\right] n^{K}n'^{K'}e^{2i\pi\alpha_{0}(n'+n)}. \end{aligned}$$

We now study the asymptotic behavior of these terms. According to Section III, we know that

$$\begin{split} \mathbf{R}_{\mathbf{e}}(n,\tau) = & \mathbb{E}\left[\mathbf{e}(n+\tau)\mathbf{e}^{*}(n)\right] = \mathbf{R}_{\mathbf{e}}(\tau),\\ & \text{and}\\ \mathbf{R}_{\mathbf{e}^{(c)}}(n,\tau) = & \mathbb{E}\left[\mathbf{e}(n+\tau)\mathbf{e}(n)\right] = \mathbf{R}_{\mathbf{e}^{(c)}}^{(2\alpha_{0})}(\tau)e^{2i\pi 2\alpha_{0}n} \end{split}$$

Using Condition 1 and well-known results on Césaro sums, we obtain after some simple manipulations that

$$\lim_{N \to \infty} \mathbf{P}_{N}(K, K') = \frac{1}{K + K' + 1} S_{\mathbf{e}}^{(0)} \left(e^{2i\pi\alpha_{0}} \right)$$
(26)

$$\lim_{N \to \infty} \mathbf{P}_{N}^{(c)}(K, K') = \frac{1}{K + K' + 1} S_{\mathbf{e}^{(c)}}^{(2\alpha_{0})} \left(e^{2i\pi\alpha_{0}} \right)$$
(27)

where $f \mapsto S_{\mathbf{e}}^{(0)}(\exp(2i\pi f))$ and $f \mapsto S_{\mathbf{e}^{(c)}}^{(2\alpha_0)}(\exp(2i\pi f))$ represent the cyclic spectrum at cyclic frequency 0 and the conjugate cyclic spectrum at cyclic frequency $2\alpha_0$ of $\mathbf{e}(n)$, respectively. Fortunately, $S_{\mathbf{e}}^{(0)}(\exp(2i\pi\alpha_0))$ and $S_{\mathbf{e}^{(c)}}^{(2\alpha_0)}(\exp(2i\pi\alpha_0))$ can be expressed more explicitly. Let $\delta \hat{\mathbf{r}}_{y^{(c)},N}^{(\alpha_0)} \coloneqq \hat{\mathbf{r}}_{y^{(c)},N}^{(\alpha_0)} - \mathbf{r}_{y^{(c)}}^{(\alpha_0)}$ denote the estimation error corresponding to the statistics $\hat{\mathbf{r}}_{y^{(c)}}^{(\alpha_0)}$. It is well known that $\sqrt{N}\delta \hat{\mathbf{r}}_{y^{(c)},N}^{(\alpha_0)}$ converges to a Gaussian distribution [20]. Let

$$\left| \mathcal{C}_{N}^{(L)} \right| \leq N^{-L/2} \sum_{n_{0},\dots,n_{L-1}=0}^{N-1} \left| \operatorname{cum}_{L} \left\{ e_{\tau_{0}}^{(\nu_{0})}(n_{0}),\dots,e_{\tau_{L-1}}^{(\nu_{L-1})}(n_{L-1}) \right\} \right|.$$
(23)

 Γ and $\Gamma^{(c)}$ denote the unconjugated/conjugated asymptotic covariance matrices

$$\begin{split} \mathbf{\Gamma} &:= \lim_{N \to \infty} N \mathbb{E} \left[\delta \hat{\mathbf{r}}_{y^{(c)},N}^{(\alpha_0)} \delta \hat{\mathbf{r}}_{y^{(c)},N}^{(\alpha_0)^*} \right] \\ \mathbf{\Gamma}^{(c)} &:= \lim_{N \to \infty} N \mathbb{E} \left[\delta \hat{\mathbf{r}}_{y^{(c)},N}^{(\alpha_0)} \delta \hat{\mathbf{r}}_{y^{(c)},N}^{(\alpha_0)^{\mathrm{T}}} \right] \end{split}$$

Since $\mathbf{y}_2(n) = \mathbf{r}_{y^{(c)}}^{(\alpha_0)} e^{2i\pi\alpha_0 n} + \mathbf{e}(n)$, it follows that

$$\sqrt{N}\delta \hat{\mathbf{r}}_{u^{(c)},N}^{(\alpha_0)} = \mathbf{E}_N(0).$$

From this, we deduce further that [19]

$$\boldsymbol{\Gamma} = \lim_{N \to \infty} \mathbb{E} \left[\mathbf{E}_{N}(0) \mathbf{E}_{N}^{*}(0) \right] = \lim_{N \to \infty} \mathbf{P}_{N}(0,0)$$

$$= S_{\mathbf{e}}^{(0)} \left(e^{2i\pi\alpha_{0}} \right)$$

$$\boldsymbol{\Gamma}^{(c)} = \lim_{N \to \infty} \mathbb{E} \left[\mathbf{E}_{N}(0) \mathbf{E}_{N}^{\mathrm{T}}(0) \right] = \lim_{N \to \infty} \mathbf{P}_{N}^{(c)}(0,0)$$

$$= S_{\mathbf{e}^{(c)}}^{(2\alpha_{0})} \left(e^{2i\pi\alpha_{0}} \right).$$
(29)

Plugging the (26)–(29) back into (25) yields

$$\lim_{N \to \infty} \mathbb{E}\left[\mathbb{E}(N)\mathbb{E}^{*}(N)\right] = \begin{bmatrix} \mathbf{\Gamma} & \frac{1}{2}\mathbf{\Gamma} & \mathbf{\Gamma}^{(c)} & \frac{1}{2}\mathbf{\Gamma}^{(c)} \\ \frac{1}{2}\mathbf{\Gamma} & \frac{1}{3}\mathbf{\Gamma} & \frac{1}{2}\mathbf{\Gamma}^{(c)} & \frac{1}{3}\mathbf{\Gamma}^{(c)} \\ \frac{1}{2}\overline{\mathbf{\Gamma}^{(c)}} & \frac{1}{2}\frac{\mathbf{\Gamma}^{(c)}}{\mathbf{\Gamma}^{(c)}} & \overline{\mathbf{\Gamma}} & \frac{1}{2}\overline{\mathbf{\Gamma}} \\ \frac{1}{2}\overline{\mathbf{\Gamma}^{(c)}} & \frac{1}{3}\overline{\mathbf{\Gamma}^{(c)}} & \frac{1}{2}\overline{\mathbf{\Gamma}} & \frac{1}{3}\overline{\mathbf{\Gamma}} \end{bmatrix}.$$

Finally, by combining the previous result and the (24), we obtain, as expected, that

$$\gamma = \frac{3}{\pi^2} \frac{\mathbf{R}_{y^{(c)}}^{(\alpha_0)^*} \mathbf{GR}_{y^{(c)}}^{(\alpha_0)}}{\left\| \mathbf{R}_{y^{(c)}}^{(\alpha_0)} \right\|^4}$$
(30)

where

$$\mathbf{R}_{y^{(c)}}^{(\alpha_0)} = \begin{bmatrix} \mathbf{r}_{y^{(c)}}^{(\alpha_0)} \\ \mathbf{r}_{y^{(c)}}^{(\alpha_0)} \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} \mathbf{\underline{\Gamma}} & -\mathbf{\underline{\Gamma}}^{(c)} \\ -\mathbf{\overline{\Gamma}}^{(c)} & \mathbf{\overline{\Gamma}} \end{bmatrix}.$$

REFERENCES

- D. Godard, "Self recovering equalization and carrier tracking in two dimensional data communications systems," *IEEE Trans. Commun.*, vol. 46, pp. 1867–1875, Nov. 1998.
- [2] 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.
- [3] L. Tong, "Joint blind signal detection and carrier recovery over fading channel," in *Proc. ICASSP*, vol. V, Detroit, MI, May 1995, pp. 1205–1208.
- [4] 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.
- [5] M. Ghogho, A. Swami, and T. Durrani, "Blind synchronization and doppler spread estimation for MSK signals in time-selective fading channels," in *Proc. ICASSP*, Istanbul, Turkey, June 2000.

- [6] —, "On blind carrier recovery in time-selective fading channels," in Proc. 31st Asilomar Conf., Pacific Grove, CA, Oct. 1999.
- [7] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase channels," *IEEE Trans. Inform. Theory*, vol. 36, pp. 312–321, Mar. 1990.
- [8] G. Zhou and G. B. Giannakis, "Harmonics in multiplicative and additive noise: performance analysis of cyclic estimators," *IEEE Trans. Signal Processing*, vol. 43, pp. 1445–1460, June 1995.
- [9] O. Besson and P. Stoica, "Frequency estimation and detection for sinusoidal signal with arbitrary envelope: A nonlinear least-square approach," in *Proc. ICASSP*, vol. 4, Seattle, WA, 1998, pp. 2209–2212.
- [10] M. Ghogho, A. Swami, and A. K. Nandi, "Nonlinear least-squares estimation for harmonics in multiplicative and additive noise," *Signal Process.*, vol. 79, no. 2, pp. 43–60, Oct. 1999.
- [11] M. Ghogho, A. Swami, and B. Garel, "Performance analysis of cyclic statistics for the estimation of harmonics in multiplicative and additive noise," *IEEE Trans. Signal Processing*, vol. 47, pp. 3235–3249, Dec. 1999.
- [12] E. Serpedin, A. Chevreuil, G. B. Giannakis, and Ph. Loubaton, "Non-data aided joint estimation of carrier frequency offset and channel using periodic modulation precoder: performance analysis," in *Proc. ICASSP*, vol. 4, Seattle, WA, 1999, pp. 2635–2638.
- [13] D. R. Brillinger, "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.
- [14] E. J. Hannan, "The nonlinear time series regression," J. Appl. Prob., vol. 8, pp. 767–780, 1972.
- [15] —, "The estimation of frequency," J. Appl. Prob., vol. 10, pp. 510–519, 1973.
- [16] 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.
- [17] P. Ciblat and P. Loubaton, "Asymptotic analysis of blind cyclic correlation based symbol-rate estimators," *IEEE Trans. Inform. Theory*, Mar. 2000, submitted for publication.
- [18] P. Ciblat, "Quelques Problèmes d'estimation Relatifs aux Télécommunications Non-Coopératives," Ph.D. dissertation (in French), Univ. Marne-la-Vallée, Marne-la-Vallée, France, 2000.
- [19] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*. New York: Springer-Verlag, 1991.
- [20] A. V. Dandawaté and G. B. Giannakis, "Asymptotic theory of mixed time averages and k th-order cyclic moment and cumulant statistics," *IEEE Trans. Inform. Theory*, vol. 41, pp. 216–232, Jan. 1995.
- [21] U. Mengali and A. D'Andrea, Synchronization Techniques for Digital Receivers. New York: Plenum, 1997.
- [22] D. R. Brillinger, *Time Series Data Analysis and Theory*. San Francisco, CA: Holden Day, 1981.
- [23] H. A. Cirpan and M. K. Tsatsanis, "Maximum likelihood blind channel estimation in the presence of frequency shifts," in *Proc. 30th Asilomar Conf. Signals, Syst., Comput.*, vol. 1, Pacific Grove, CA, Nov. 1996, pp. 713–717.
- [24] K. E. Scott and E. B. Olasz, "Simultaneous clock phase and frequency offset estimation," *IEEE Trans. Commun.*, vol. 43, pp. 2263–2270, July 1995.
- [25] R. Mehlan, Y. Chen, and H. Meyr, "A fully digital feedforward MSK demodulator with joint frequency offset and symbol timing estimation for burst mode mobile radio," *IEEE Trans. Veh. Technol.*, vol. 42, pp. 434–443, Nov. 1993.
- [26] M. Morelli and U. Mengali, "Joint frequency and timing recovery for MSK-type modulation," *IEEE Trans. Commun.*, vol. 47, pp. 938–946, June 1999.

Philippe Ciblat was born in Paris, France, in 1973. He received the Eng. degree from Ecole Nationale Supérieure des Télécommunications (ENST), Paris, the M.Sc. in signal processing from University of Paris-Sud, Orsay, France, in 1996, and the Ph.D. degree from the University of Marne-la-Vallée, Champs sur Marne, France, in July 2000.

From October 2000 to June 2001, he was a postdoctoral researcher with the Communications and Remote Sensing Department, Université Catholique de Louvain, Louvain-la-Neuve, Belgium. He is currently Associate Professor with the Département Communications et Electronique, ENST. His research areas include statistical and digital signal processing, especially blind equalization and frequency estimation.

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

From 1982 to 1986, he was Member of Technical Staff of 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 ENST, where he worked in the Signal Processing Department. Since 1995, he has been Professor of Electrical Engineering at Marne la Vallée University, Champs sur Marne, France. From 1996 to 2000, he was Director of the Laboratoire Système de Communication, Marne la Vallée University, and is now a member of the Laboratoire Traitement et Communication de l'Information, CNRS/ENST. His present research interests are in statistical signal processing and digital communications with a special emphasis on blind equalization, multiuser communication systems, and multicarrier modulations.

Dr. Loubaton is currently Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and IEEE COMMUNICATIONS LETTERS and is a member of the IEEE Signal Processing for Communications Technical Committee.

Erchin Serpedin received (with highest distinction) the Dipl. E.E. degree from the Polytechnic Institute of Bucharest, Bucharest, Romania, in 1991. He received the specialization degree in signal processing and transmission of information from Ecole Supérieure d'Electricité, 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 and between January and June 1999, he was a Lecturer with the University of Virginia. In July 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.

Dr. Serpedin received the NSF Career Award in November 2000.



Georgios B. Giannakis (F'97) received the Dipl.Eng. degree 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, Charlottesville, in 1987,

where he became a Professor of electrical engineering in 1997. Since 1999, he has been with the University of Minnesota, Minneapolis, as a Professor of electrical and computer engineering. His general interests span the areas of communications and signal processing, estimation and detection theory, time-series analysis and system identification, subjects on which he has published more than 120 journal papers, 250 conference papers, and two edited books. Current research topics focus on transmitter and receiver diversity techniques for single- and multiuser fading communication channels, redundant precoding and space-time coding for block transmissions, and multicarrier and wideband wireless communication systems. He is a frequent consultant to the telecommunications industry.

Dr. Giannakis was the (co-) recipient of best paper awards from the IEEE Signal Processing (SP) Society in 1992, 1998, and 2000. He also received the Society's Technical Achievement Award in 2000. He co-organized the IEEE-SP Workshops on HOS in 1993, SSAP in 1996, and SPAWC in 1997 and guest (co-) edited four special issues. He has served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the the IEEE SIGNAL PROCESSING LETTERS, a secretary of the SP Conference Board, a member of the SP Publications Board, and a member and vice chair of the Statistical Signal and Array Processing Committee. He is a member of the Editorial Board for the PROCEEDINGS OF THE IEEE, he chairs the SP for Communications Technical Committee, and serves as the Editor in Chief for the IEEE SIGNAL PROCESSING LETTERS. He is a member of the IEEE Fellows Election Committee and the IEEE-SP Society's Board of Governors.