Asymptotic Analysis of Blind Cyclic Correlation-Based Symbol-Rate Estimators

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

***Abstract*—This paper considers the problem of blind symbol rate estimation of signals linearly modulated by a sequence of unknown symbols. Oversampling the received signal generates cyclostationary statistics that are exploited to devise symbol-rate estimators by maximizing in the cyclic domain a (possibly weighted) sum of modulus squares of cyclic correlation estimates. Although quite natural, the asymptotic (large sample) performance of this estimator has not been studied rigorously. The consistency and asymptotic normality of this symbol-rate estimator is established when the number of samples N converges to infinity. It is shown that this estimator exhibits a fast convergence rate (proportional to $N^{-3/2}$), and it admits a simple closed-form expression for its asymptotic variance. This asymptotic expression enables performance analysis of the rate estimator as a function of the number of estimated cyclic correlation coefficients and the weighting matrix. A justification for the high performance of the unweighted estimator in high signal-to-noise scenarios is also provided.**

***Index Terms*—Cumulant, cyclostationary, estimation, frequency, spectrum estimation, symbol rate.**

I. Introduction

LET $y_a(t)$ denote the complex envelope of the continuous-time[1] received signal, which is supposed to be transmitted by an *unknown* communication source that employs linear digital modulation. Signal $y_a(t)$ can thus be expressed as

$$y_a(t) = e^{2i\pi\Delta f_0 t} \sum_{k\in\mathbb{Z}} s_k h_a(t - kT_s) + w_a(t) \tag{1}$$

where $\{s_n\}_{n\in\mathbb{Z}}$ is a zero-mean unit variance independent and identically distributed (i.i.d.) sequence of symbols, $1/T_s$ stands for the baud rate of the transmitter, $h_a(t)$ denotes the convolution of the shaping filter with the unknown multipath channel, Δf_0 stands for the carrier frequency offset, and $w_a(t)$ represents a normally distributed noise. In certain applications, estimation of the unknown parameters of the received waveform (the symbol-rate $1/T_s$, the carrier frequency/phase offset Δf_0, and the symbol alphabet) has to be performed blindly from the received waveform, without knowledge of the transmitted symbols. Potential applications include passive listening, automatic classification of modulation, and blind synchronization of high-speed distributed networks, where the receiver has to be synchronized without transmitting a pilot/training sequence, for bandwidth efficiency reasons.

The goal of this paper is blind estimation of the symbol rate $1/T_s$ from a sampled version $y(n) = y_a(nT_e)$ of continuous-time received waveform $y_a(t)$, where the sampling period T_e is taken sufficiently small to satisfy $T_e < T_s/4$. The condition $T_e < T_s/4$, although not restrictive from a practical standpoint, is adopted because it enables a relatively simple closed-form and "interpretable" expression for the asymptotic variance of the symbol-rate estimator. For the sake of simplicity, it is also assumed that $\Delta f_0 = 0$, i.e., no carrier frequency offset is present. The reader may check that this assumption is not restrictive and that all the following results remain true if $\Delta f_0 \neq 0$.

As in every parameter estimation problem, maximum likelihood of T_s can be formulated, and implemented using an expection–maximization (EM) algorithm. However, in our particular context, several nuisance parameters such as the channel impulse response or the transmitted symbols have to be jointly estimated in the maximum-likelihood sense. The resulting EM algorithm is therefore in practice difficult to use. We rather focus on a classical suboptimum estimator of T_s which relies on the observation that $y_a(t)$ is a cyclostationary signal and its cyclic frequencies are integer multiples of $1/T_s$. The bandwidth of $y_a(t)$ is assumed to be the interval $[-(1+\rho)/(2T_s), (1+\rho)/(2T_s)]$, where the parameter ρ represents the excess bandwidth and belongs to the interval (0, 1]. It follows that $-1/T_s$, 0, and $1/T_s$ are the sole nonnegligible cyclic frequencies. The discrete-time signal $y(n)$ is thus cyclostationary and the parameter α_0 defined by

$$\alpha_0 := \frac{T_e}{T_s}$$

represents its unique strictly positive cyclic frequency. Note that the condition: $T_e < T_s/4$ implies that $\alpha_0 < 1/4$. Estimating T_s is thus equivalent to estimating α_0. For this, we introduce the autocovariance function of $y(n)$, denoted by $r_y(n, \tau)$ and defined by $r_y(n, \tau) := \mathbb{E}[y(n+\tau)y^*(n)]$. The superscript $*$ stands for Hermitian conjugate. We obtain that

$$r_y(n, \tau) = \sum_{k=-1}^{1} r_y^{(k\alpha_0)}(\tau) e^{2i\pi k\alpha_0 n} \tag{2}$$

Manuscript received April 2, 2000; revised February 7, 2002. This work was supported by a DGA/CNRS Fellowship.

P. Ciblat is with the Département Communications et Electronique, Ecole Nationale Supérieure des Télécommunications, 75634 Paris Cedex 13, France (e-mail: philippe.ciblat@enst.fr).

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

Communicated by J. A. O'Sullivan, Associate Editor for Detection and Estimation.

Publisher Item Identifier S 0018-9448(02)05336-1.

[1]The subscript a is used to denote continuous-time analog signals.

where $r_y^{(k\alpha_0)}(\tau)$ is the cyclic correlation of $y(n)$ at cyclic frequency $k\alpha_0$ and lag τ. Let

$$\boldsymbol{r}_y(n) := [r_y(n, -\Upsilon), \ldots, r_y(n, \Upsilon)]^T$$

denote the vector of $2\Upsilon + 1$ autocovariance coefficients, where the superscript T stands for transposition. It follows that

$$\boldsymbol{r}_y(n) = \sum_{k=-1}^{1} \boldsymbol{r}_y^{(k\alpha_0)} e^{2i\pi k\alpha_0 n}, \tag{3}$$

with the $(2\Upsilon+1)\times 1$ vector $\boldsymbol{r}_y^{(\alpha)} := [r_y^{(\alpha)}(-\Upsilon), \ldots, r_y^{(\alpha)}(\Upsilon)]^T$. Since, for each τ, $r_y^{(\alpha)}(\tau) = 0$ when α is different from $-\alpha_0$, 0, and α_0, it is clear that, $\boldsymbol{r}_y^{(\alpha)} = 0$ when $\alpha \neq -\alpha_0, 0, \alpha_0$. Therefore, we can consider the following classical symbol rate estimator (see, e.g., [6], where similar criterion is considered)

$$\alpha_0 := \arg\max_{\alpha\in\mathcal{I}} J_{\boldsymbol{W}}(\alpha)$$

where

$$J_{\boldsymbol{W}}(\alpha) := \boldsymbol{r}_y^{(\alpha)^*} \boldsymbol{W} \boldsymbol{r}_y^{(\alpha)}$$

and where $\mathcal{I}$ denotes a closed interval included in $(0, 1/4)$ and the Hermitian matrix $\boldsymbol{W}$ is positive-definite ($\boldsymbol{W} = \boldsymbol{W}^* \geq \boldsymbol{0}$). Selection of the matrix $\boldsymbol{W}$ will be addressed later. In practice, for a given cyclic frequency α, $\boldsymbol{r}_y^{(\alpha)}$ has to be estimated from a finite number N of samples $\{y(n)\}_{n=0,\ldots,N-1}$. The standard sample estimate of $\boldsymbol{r}_N^{(\alpha)}$ is given by

$$\hat{\boldsymbol{r}}_N^{(\alpha)} := \frac{1}{N} \sum_{n=0}^{N-1} \boldsymbol{y}_2(n) e^{-2i\pi\alpha n}$$

where

$$\boldsymbol{y}_2(n) := [y(n-\Upsilon)y^*(n), \ldots, y(n+\Upsilon)y^*(n)]^T.$$

The cyclic frequency α_0 can thus be estimated by the element $\hat{\alpha}_{N,\boldsymbol{W}}$ of $\mathcal{I}$ defined by

$$\hat{\alpha}_{N,\boldsymbol{W}} := \arg\max_{\alpha\in\mathcal{I}} J_{N,\boldsymbol{W}}(\alpha) \tag{4}$$

where $J_{N,\boldsymbol{W}}(\alpha)$ is the sampled version of $J_{\boldsymbol{W}}(\alpha)$

$$J_{N,\boldsymbol{W}}(\alpha) = \hat{\boldsymbol{r}}_N^{(\alpha)^*} \boldsymbol{W} \hat{\boldsymbol{r}}_N^{(\alpha)}. \tag{5}$$

In this paper, we prove consistency and asymptotic normality of $\hat{\alpha}_{N,\boldsymbol{W}}$. We also show that the convergence rate of $\hat{\alpha}_{N,\boldsymbol{W}}$ is $N^{-3/2}$ and calculate in closed form its asymptotic variance defined by

$$\gamma_{\boldsymbol{W}} = \lim_{N\to\infty} N^3 \mathbb{E}\left[(\hat{\alpha}_{N,\boldsymbol{W}} - \alpha_0)^2\right].$$

This expression enables performance analysis of $\hat{\alpha}_{N,\boldsymbol{W}}$ as a function of $\boldsymbol{W}$ and the number $2\Upsilon + 1$ of the estimated cyclic correlation coefficients used.

Statistical analysis of cyclostationary stochastic processes was studied by several authors (see, e.g., [11], [8], and the references therein). However, most of these works addressed cyclic correlation and spectrum estimation problems. Estimation of cyclic frequencies has been relatively less popular. Some previous works addressed the problem of testing if a given frequency α is a cyclic frequency or not (see, e.g., [6], [9], where the cost function $J_{N,\boldsymbol{W}}(\alpha)$ and higher order statistical tests are considered). However, thorough analysis of the asymptotic performance of the estimate $\hat{\alpha}_{N,\boldsymbol{W}}$ has not yet been considered. Our work relates also to various papers devoted to frequency estimation in multiplicative and additive noise [2], [10], [23]. In these works, the observation $y(n)$ is modeled as

$$y(n) = \lambda(n)e^{2i\pi n f_0} + w(n)$$

where $w(n)$ denotes additive white Gaussian noise, and the "multiplicative" noise $\lambda(n)$ is a real and noncircular stationary stochastic process. In [10] and [23], the observation is real, so that $\exp(2i\pi n f_0)$ is replaced by $\cos(2\pi n f_0)$. Using the observation that $2f_0$ is the unique conjugate cyclic frequency of $y(n)$ (i.e., $\mathbb{E}[y^2(n)] = r_y^{(2\alpha_0)}(\tau)\exp(2i\pi n 2f_0)$), the authors proposed to estimate $2f_0$ by maximizing with respect to α the criterion

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

which is similar to our cost function in (5) when $\Upsilon = 0$.

The starting point of our work is based on the observation, used in [2], [10], [20], [21], and [23], that certain cyclic frequency estimation problems can be formulated as the estimation of a number of sinusoids embedded in additive noise and the cost function $J_{N,\boldsymbol{I}}(\alpha)$ is equivalent to a periodogram ($\boldsymbol{I}$ stands for the identity matrix). The most popular approach to study the asymptotic behavior of periodogram estimates is to introduce an auxiliary nonlinear least squares (NLS) problem [2], [4], [10], [12], [14], [20], [21], and [23].

However, this approach cannot be used to analyze $\hat{\alpha}_{N,\boldsymbol{W}}$ for $\boldsymbol{W} \neq \boldsymbol{I}$. More important, unless the number of cyclic correlation coefficients taken into account is reduced to one as in [2], [10], [20], [21], [23] (i.e., $\Upsilon = 0$), calculating the variance of $\hat{\alpha}_{N,\boldsymbol{I}}$ by the NLS approach necessitates very complicated computations which do not lead to interpretable closed-form expressions. This is a major result of this paper because, as it will be shown later, the performance of the estimate can be enhanced by choosing appropriately the parameter Υ. In this paper, we show that the use of the auxiliary NLS criterion is not necessary, and that the asymptotic properties of $\hat{\alpha}_{N,\boldsymbol{W}}$ can be established by using a much simpler alternative approach, obtained by generalizing the analysis of [13].

This paper is organized as follows. In Section II, we explain the connections between the estimation of α_0 and the estimation of the frequency of a sinusoid corrupted by additive noise. In Section III, we state our main results and derive a closed-form expression for the asymptotic variance of $\hat{\alpha}_{N,\boldsymbol{W}}$. In Section IV, we rely on this expression to discuss the selection of the weighting matrix $\boldsymbol{W}$ and the number of cyclic correlations $2\Upsilon + 1$ that minimize the asymptotic variance of $\hat{\alpha}_{N,\boldsymbol{W}}$. We indicate that the choice $\boldsymbol{W} = \boldsymbol{I}$ appears quite appropriate. We also show that the selection of parameter Υ plays an important role in improving the performance of the symbol rate estimator. By choosing Υ greater than the memory of the channel $h(z) = \sum_k h_a(kT_e)z^{-k}$ leads to improved performance of the estimator. In Section V, we finally illustrate the performance of $\hat{\alpha}_{N,\boldsymbol{W}}$ by means of numerical evaluations.

II. Harmonic Retrieval Links

Estimating the cycle α_0 can be reduced to estimating frequencies of a number of sinusoids embedded in noise [23], [10], [20], [21]. To show this equivalence, let $\boldsymbol{e}_2(n) = [e_{-\Upsilon}(n), \ldots, e_{\Upsilon}(n)]^T$ be the mean-compensated $(2\Upsilon+1)$-dimensional stochastic process defined by

$$\boldsymbol{e}_2(n) := \boldsymbol{y}_2(n) - \mathbb{E}[\boldsymbol{y}_2(n)] \tag{6}$$

where $\mathbb{E}[\boldsymbol{y}_2(n)] := \boldsymbol{r}_y(n)$. According to (3) and (6), it follows readily that

$$\boldsymbol{y}_2(n) = \sum_{k=-1}^{1} \boldsymbol{r}_y^{(k\alpha_0)} e^{2i\pi k\alpha_0 n} + \boldsymbol{e}_2(n). \tag{7}$$

Thus, $\boldsymbol{y}_2(n)$ can be interpreted as a sum of vector-valued complex frequencies 0, α_0, and $-\alpha_0$ corrupted by the zero-mean additive noise $\boldsymbol{e}_2(n)$. Moreover, the criterion $J_{N,\boldsymbol{W}}(\alpha)$ in (5) is simply a weighted periodogram because

$$J_{N,\boldsymbol{W}}(\alpha) = \left\| \frac{1}{N} \sum_{n=0}^{N-1} \boldsymbol{y}_2(n) e^{-2i\pi\alpha n} \right\|_{\boldsymbol{W}}^2$$

with the norm $\|\boldsymbol{x}\|_{\boldsymbol{W}}^2 = \boldsymbol{x}^*\boldsymbol{W}\boldsymbol{x}$ for any $(2\Upsilon+1)$-dimensional vector $\boldsymbol{x}$.

Several works have been devoted to frequency estimation [4], [12]–[14]. However, the differences between the present context and these works are: i) $\boldsymbol{y}_2(n)$ is multivariate, ii) the criterion is weighted by the matrix $\boldsymbol{W}$, and iii) $\boldsymbol{e}_2(n)$ is not stationary but cyclostationary.

However, [20] and [21] showed that the approach of [4], [12], [14] can be generalized when $\Upsilon = 0$ and $\boldsymbol{e}_2(n)$ is periodically correlated. In this context, one can apparently follow the classical approach [4], [12], [14], which is based on the NLS estimation auxiliary problem

$$\left[\hat{\boldsymbol{\theta}}_N, \hat{\alpha}_N^{(K)}\right] := \arg \min_{\alpha\in\mathcal{I}, \boldsymbol{\theta}\in\mathbb{C}^3} K_N(\boldsymbol{\theta}, \alpha)$$

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

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

with $\boldsymbol{\theta}^T := [\theta_{-1}, \theta_0, \theta_1]$ (recall that $\Upsilon = 0$ in [20] and [21]). 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 $\hat{\alpha}_{N,\boldsymbol{I}}$ are asymptotically equivalent. In [20] and [21], evaluation of the asymptotic variance of $\hat{\alpha}_{N,\boldsymbol{I}}$ is performed by calculating the asymptotic variance of the vector $[\hat{\boldsymbol{\theta}}_N^T, \hat{\alpha}_N^{(K)T}]$. This is achieved by using a second-order Taylor expansion of $K_N(\boldsymbol{\theta}, \alpha)$ around the true values of the four parameters, which provides the asymptotic covariance matrix of the four-dimensional vector $[\hat{\boldsymbol{\theta}}_N^T, \hat{\alpha}_N^{(K)T}]$. This approach is also used in [2], [10], and [23], in the context of frequency estimation in presence of multiplicative and additive noise.

It can be shown that the NLS approach can be generalized to the case when $\boldsymbol{e}_2(n)$ is a general nonstationary process. However, the NLS approach appears more difficult to extend when $\Upsilon > 0$. In this case, $\boldsymbol{\theta}$ is a $3(2\Upsilon+1)$-dimensional vector, and the calculations needed to evaluate the covariance matrix of $[\hat{\boldsymbol{\theta}}_N^T, \hat{\alpha}_N^{(K)T}]$ do not lead to a simple and interpretable expression for the asymptotic covariance matrix of $\hat{\alpha}_N^{(K)}$. Moreover, the NLS approach cannot be generalized to analyze the properties of $\hat{\alpha}_{N,\boldsymbol{W}}$ for $\boldsymbol{W} \neq \boldsymbol{I}$. In the next section, the asymptotic performance analysis of $\hat{\alpha}_{N,\boldsymbol{W}}$ is performed by generalizing the approach of [13], which is based on a direct study of the argument that maximizes the periodogram.

III. Asymptotic Performance

In the following, the overbar $\overline{}$ will be used to denote complex conjugation. If $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L$ are random vectors, the notation $\operatorname{cum}_L(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L)$ will stand for the Lth-order cumulant tensor of vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L$. The following mixing condition will be required to establish the asymptotic performance of the symbol-rate estimator.

Assumption 1: Let $\boldsymbol{e}_2(n)$ be the multivariate random process defined by (6) and define $\boldsymbol{e}_2^{(0)}(n) := \boldsymbol{e}_2(n)$ and $\boldsymbol{e}_2^{(1)}(n) := \overline{\boldsymbol{e}}_2(n)$. We assume that $\boldsymbol{e}_2(n)$ satisfies the following mixing condition:

$$\forall L, \exists \mathcal{M}_L, \forall n_1, \forall(\nu_1, \ldots, \nu_L) \in \{0, 1\}^L,$$
$$\sum_{n_2, \ldots, n_L=-\infty}^{+\infty} \left\| \operatorname{cum}_L\left(\boldsymbol{e}_2^{(\nu_1)}(n_1), \ldots, \boldsymbol{e}_2^{(\nu_L)}(n_L)\right) \right\| \le \mathcal{M}_L. \tag{8}$$

The mixing condition (8) is not restrictive and it is satisfied when the impulse response of the filter $h_a(t)$ has finite memory. Under Assumption 1, it is possible to prove the following lemma which generalizes the scalar counterpart in [4], [12]–[14], [20], and [21].

Lemma 1: Assume that $\boldsymbol{e}_2(n)$ satisfies Assumption 1 and define

$$\boldsymbol{s}_N^{(K)}(\alpha) := \frac{1}{N^{K+1}} \sum_{n=0}^{N-1} n^K \boldsymbol{e}_2(n) e^{2i\pi\alpha n}. \tag{9}$$

Then

$$\forall K \in \mathbb{N}, \quad \sup_{\alpha\in[0,1]} \left\| \boldsymbol{s}_N^{(K)}(\alpha) \right\| \xrightarrow{\text{a.s.}} 0, \qquad \text{as } N \to \infty. \tag{10}$$

Proof: See Appendix I.

A. Consistency

In this subsection, we study the consistency of $\hat{\alpha}_{N,\boldsymbol{W}}$, and establish that $N(\hat{\alpha}_{N,\boldsymbol{W}} - \alpha_0)$ converges to zero almost surely. This property is quite useful in establishing the asymptotic normality of $\hat{\alpha}_{N,\boldsymbol{W}}$. The following result holds.

Theorem 1: Assume that the positive matrix $\boldsymbol{W}$ satisfies $\boldsymbol{r}_y^{(\alpha_0)^*}\boldsymbol{W}\boldsymbol{r}_y^{(\alpha_0)} > 0$ (if $\boldsymbol{W} > 0$, this condition is automatically satisfied). Then, as $N \to \infty$

$$\hat{\alpha}_{N,\boldsymbol{W}} - \alpha_0 \to 0 \quad \text{and} \quad N(\hat{\alpha}_{N,\boldsymbol{W}} - \alpha_0) \to 0$$

almost surely.

Proof: We first note that the sequence $\{\hat{\alpha}_{N,\boldsymbol{W}}\}_{N\geq 0}$ belongs to the compact set $\mathcal{I} \subset (0, 1/4)$. Therefore, in order to prove that $\hat{\alpha}_{N,\boldsymbol{W}}$ converges almost surely to α_0, it is sufficient to establish that every convergent subsequence extracted from $\hat{\alpha}_{N,\boldsymbol{W}}$ converges to α_0. For this, we consider a subsequence $\{\hat{\alpha}_{\phi(N),\boldsymbol{W}}\}_{N\in\mathbb{N}}$ converging to a certain value $\alpha_1 \in \mathcal{I}$, and we prove that $\alpha_1 = \alpha_0$.

Since $\hat{\alpha}_{N,\boldsymbol{W}}$ maximizes $J_{N,\boldsymbol{W}}(\alpha)$, it follows that

$$J_{N,\boldsymbol{W}}(\hat{\alpha}_{N,\boldsymbol{W}}) \geq J_{N,\boldsymbol{W}}(\alpha_0).$$

This inequality remains true for the subsequence, and we obtain immediately that

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

exists almost surely and is nonnegative.

Using the expression (7) of $\boldsymbol{y}_2(n)$, one observes that $J_{N,\boldsymbol{W}}(\alpha)$ can be decomposed into four terms

$$J_{N,\boldsymbol{W}}(\alpha) = \boldsymbol{t}_N^*(\alpha)\boldsymbol{W}\boldsymbol{t}_N(\alpha) + \boldsymbol{s}_N^{(0)*}(\alpha)\boldsymbol{W}\boldsymbol{t}_N(\alpha) + \boldsymbol{t}_N^*(\alpha)\boldsymbol{W}\boldsymbol{s}_N^{(0)}(\alpha) + \boldsymbol{s}_N^{(0)*}(\alpha)\boldsymbol{W}\boldsymbol{s}_N^{(0)}(\alpha)$$

where

$$\boldsymbol{t}_N(\alpha) := \frac{1}{N}\sum_{n=0}^{N-1}\boldsymbol{r}_y(n)\exp(-2i\pi\alpha n)$$

and $\boldsymbol{s}_N^{(0)}(\alpha)$ is defined by (9) for $K = 0$. It is easy to verify that $\boldsymbol{t}_N(\alpha)$ is bounded in N and α. Using Lemma 1, it follows that

$$\lim_{N\to\infty} J_{\phi(N),\boldsymbol{W}}(\alpha_0) \stackrel{\text{a.s.}}{=} \lim_{N\to\infty}\left\|\boldsymbol{t}_{\phi(N)}(\alpha_0)\right\|_{\boldsymbol{W}}^2 \tag{11}$$

and

$$\lim_{N\to\infty} J_{\phi(N),\boldsymbol{W}}\left(\hat{\alpha}_{\phi(N),\boldsymbol{W}}\right) \stackrel{\text{a.s.}}{=} \lim_{N\to\infty}\left\|\boldsymbol{t}_{\phi(N)}\left(\hat{\alpha}_{\phi(N),\boldsymbol{W}}\right)\right\|_{\boldsymbol{W}}^2. \tag{12}$$

Recalling the expansion

$$\boldsymbol{r}_y(n) = \sum_{k=-1}^{1}\boldsymbol{r}_y^{(k\alpha_0)}\exp(2i\pi k\alpha_0 n)$$

it follows that

$$\lim_{N\to\infty} J_{\phi(N),\boldsymbol{W}}(\alpha_0) \stackrel{\text{a.s.}}{=} \left\|\boldsymbol{r}_y^{(\alpha_0)}\right\|_{\boldsymbol{W}}^2.$$

In order to evaluate the limit (12), we need to introduce the following simple lemma from [14].

Lemma 2: 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

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

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

$$\begin{aligned} q_N(c_N) &\to 0, && \text{if } c \neq 0 \\ q_N(c_N) &\to 0, && \text{if } c = 0 \text{ and } N|c_N - c| \to \infty \\ q_N(c_N) &\to e^{i\beta}\mathrm{sinc}(\beta), && \text{if } c = 0 \text{ and } \\ & && d_N = N(c_N - c) \to \beta \in \mathbb{R}. \end{aligned}$$

Equation (3) implies that

$$\lim_{N\to\infty} J_{\phi(N),\boldsymbol{W}}\left(\hat{\alpha}_{\phi(N),\boldsymbol{W}}\right) \stackrel{\text{a.s.}}{=} \lim_{N\to\infty}\left\|\sum_{k=-1}^{1}\boldsymbol{r}_y^{(k\alpha_0)}q_{\phi(N)}\left(k\alpha_0 - \hat{\alpha}_{\phi(N),\boldsymbol{W}}\right)\right\|_{\boldsymbol{W}}^2.$$

As α_1 and α_0 belong to $\mathcal{I} \subset (0, 1/4)$, the limits of

$$(-\alpha_0 - \hat{\alpha}_{\phi(N),\boldsymbol{W}}) \quad \text{and} \quad -\hat{\alpha}_{\phi(N),\boldsymbol{W}}, \qquad \text{as } N \to \infty$$

are not equal to zero (modulo 1). Lemma 2 yields that

$$\lim_{N\to\infty} J_{\phi(N),\boldsymbol{W}}\left(\hat{\alpha}_{\phi(N),\boldsymbol{W}}\right) \stackrel{\text{a.s.}}{=} \lim_{N\to\infty}\left\|\boldsymbol{r}_y^{(\alpha_0)}q_{\phi(N)}\left(\alpha_0 - \hat{\alpha}_{\phi(N),\boldsymbol{W}}\right)\right\|_{\boldsymbol{W}}^2. \tag{13}$$

If $\alpha_1 \neq \alpha_0$, Lemma 2 implies that $\Delta J \stackrel{\text{a.s.}}{=} -\|\boldsymbol{r}_y^{(\alpha_0)}\|_{\boldsymbol{W}}^2 < 0$, which contradicts the condition $\Delta J \geq 0$. Therefore, $\alpha_1 = \alpha_0$ and the first part of Theorem 1 is established.

In order to prove the second part, we consider the sequence $\{b_N\}_{N\in\mathbb{N}}$ defined by $b_N := N(\hat{\alpha}_{N,\boldsymbol{W}} - \alpha_0)$. 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 2, the corresponding ΔJ is equal to $-\|\boldsymbol{r}_y^{(\alpha_0)}\|_{\boldsymbol{W}}^2$, 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 2, we obtain that $\Delta J = \|\boldsymbol{r}_y^{(\alpha_0)}\|_{\boldsymbol{W}}^2(\mathrm{sinc}^2(\beta) - 1)$. As $\Delta J \geq 0$, β must be equal to 0. This proves that

$$N(\hat{\alpha}_{N,\boldsymbol{W}} - \alpha_0) \xrightarrow{\text{a.s.}} 0, \qquad \text{as } N \to \infty.$$

B. Asymptotic Normality of the Estimate

The asymptotic normality and convergence rate of $\hat{\alpha}_{N,\boldsymbol{W}}$ are next obtained using a second-order Taylor expansion of $J_{N,\boldsymbol{W}}(\alpha)$ around α_0

$$\left.\frac{\partial J_{N,\boldsymbol{W}}(\alpha)}{\partial\alpha}\right|_{\alpha=\hat{\alpha}_{N,\boldsymbol{W}}} - \left.\frac{\partial J_{N,\boldsymbol{W}}(\alpha)}{\partial\alpha}\right|_{\alpha=\alpha_0} = \left.\frac{\partial^2 J_{N,\boldsymbol{W}}(\alpha)}{\partial\alpha^2}\right|_{\alpha=\tilde{\alpha}_{N,\boldsymbol{W}}}\delta\hat{\alpha}_{N,\boldsymbol{W}} \tag{14}$$

where the estimation error $\delta\hat{\alpha}_{N,\boldsymbol{W}} := \hat{\alpha}_{N,\boldsymbol{W}} - \alpha_0$ and the scalar $\tilde{\alpha}_{N,\boldsymbol{W}}$ belongs to $[\alpha_0, \hat{\alpha}_{N,\boldsymbol{W}}]$ (or $[\hat{\alpha}_{N,\boldsymbol{W}}, \alpha_0]$). Taking into account that $\partial J_{N,\boldsymbol{W}}(\alpha)/\partial\alpha|_{\hat{\alpha}_{N,\boldsymbol{W}}} = 0$, it follows from (14) that

$$\delta\hat{\alpha}_{N,\boldsymbol{W}} = -\left[\left.\frac{\partial^2 J_{N,\boldsymbol{W}}(\alpha)}{\partial\alpha^2}\right|_{\alpha=\tilde{\alpha}_{N,\boldsymbol{W}}}\right]^{-1}\left.\frac{\partial J_{N,\boldsymbol{W}}(\alpha)}{\partial\alpha}\right|_{\alpha=\alpha_0}. \tag{15}$$

Define

$$\mathcal{A}_{N,\boldsymbol{W}} := \frac{1}{N^2}\left.\frac{\partial^2 J_{N,\boldsymbol{W}}(\alpha)}{\partial\alpha^2}\right|_{\alpha=\tilde{\alpha}_{N,\boldsymbol{W}}} \tag{16}$$

$$\mathcal{B}_{N,\boldsymbol{W}} := \frac{1}{\sqrt{N}}\left.\frac{\partial J_{N,\boldsymbol{W}}(\alpha)}{\partial\alpha}\right|_{\alpha=\alpha_0}. \tag{17}$$

Plugging (16) and (17) back into (15), we obtain

$$N^{\frac{3}{2}}\,\delta\hat{\alpha}_{N,\boldsymbol{W}} = -\mathcal{A}_{N,\boldsymbol{W}}^{-1}\,\mathcal{B}_{N,\boldsymbol{W}}. \tag{18}$$

In order to analyze the asymptotic properties of $N^{\frac{3}{2}}\delta\hat{\alpha}_{N,\boldsymbol{W}}$, we have to study the asymptotic behavior of $\mathcal{A}_{N,\boldsymbol{W}}$ and $\mathcal{B}_{N,\boldsymbol{W}}$.

1) Asymptotic Behavior of $\mathcal{A}_{N,\boldsymbol{W}}$: After some simple algebraic manipulations, it can be checked that $\mathcal{A}_{N,\boldsymbol{W}}$ can be expressed in terms of $(1/N^K)\partial^K\hat{\boldsymbol{r}}_N^{(\alpha)}/\partial\alpha^K|_{\alpha=\tilde{\alpha}_{N,\boldsymbol{W}}}$ with $K \in \{0, 1, 2\}$. It is also easy to check that the sequence $\{\tilde{\alpha}_{N,\boldsymbol{W}}\}_{N\in\mathbb{N}}$ satisfies Theorem 1. Using Lemmas 1 and 2, it is straightforward to show the following.

Lemma 3:

$$\frac{1}{N^K}\left.\frac{\partial^K\hat{\boldsymbol{r}}_N^{(\alpha)}}{\partial\alpha^K}\right|_{\alpha=\tilde{\alpha}_{N,\boldsymbol{W}}} \to \frac{(-2i\pi)^K}{K+1}\,\boldsymbol{r}_y^{(\alpha_0)} \quad \text{almost surely,} \quad \text{as } N\to\infty.$$

From this, we immediately obtain the following.

Proposition 1:

$$\mathcal{A}_{N,\boldsymbol{W}} \to \gamma_{\mathcal{A}} \quad \text{almost surely, as } N\to\infty, \quad \text{with} \quad \gamma_{\mathcal{A}} := -\frac{2\pi^2}{3}\,\boldsymbol{r}_y^{(\alpha_0)^*}\boldsymbol{W}\boldsymbol{r}_y^{(\alpha_0)}.$$

2) Asymptotic Behavior of $\mathcal{B}_{N,\boldsymbol{W}}$: Using (6), it follows that $\mathcal{B}_{N,\boldsymbol{W}}$ can be decomposed into the following three terms:

$$\mathcal{B}_{N,\boldsymbol{W}} = \mathcal{B}_{N,\boldsymbol{W}}^{(1)} + \mathcal{B}_{N,\boldsymbol{W}}^{(2)} + \mathcal{B}_{N,\boldsymbol{W}}^{(3)} \tag{19}$$

where

$$\begin{aligned}
\mathcal{B}_{N,\boldsymbol{W}}^{(1)} &:= \frac{-4\pi}{N^2\sqrt{N}}\mathfrak{Im}\left[\left(\sum_{n=0}^{N-1}\boldsymbol{r}_y^*(n)ne^{2i\pi\alpha_0 n}\right)\boldsymbol{W} \times\left(\sum_{n=0}^{N-1}\boldsymbol{r}_y(n)e^{-2i\pi\alpha_0 n}\right)\right]\\
\mathcal{B}_{N,\boldsymbol{W}}^{(2)} &:= \frac{-4\pi}{N^2\sqrt{N}}\mathfrak{Im}\left[\left(\sum_{n=0}^{N-1}\boldsymbol{r}_y^*(n)ne^{2i\pi\alpha_0 n}\right)\boldsymbol{W} \times\left(\sum_{n=0}^{N-1}\boldsymbol{e}_2(n)e^{-2i\pi\alpha_0 n}\right)\right.\\
&\qquad\left.+\left(\sum_{n=0}^{N-1}\boldsymbol{e}_2^*(n)ne^{2i\pi\alpha_0 n}\right)\boldsymbol{W} \times\left(\sum_{n=0}^{N-1}\boldsymbol{r}_y(n)e^{-2i\pi\alpha_0 n}\right)\right]\\
\mathcal{B}_{N,\boldsymbol{W}}^{(3)} &:= \frac{-4\pi}{N^2\sqrt{N}}\mathfrak{Im}\left[\left(\sum_{n=0}^{N-1}\boldsymbol{e}_2^*(n)ne^{2i\pi\alpha_0 n}\right)\boldsymbol{W} \times\left(\sum_{n=0}^{N-1}\boldsymbol{e}_2(n)e^{-2i\pi\alpha_0 n}\right)\right]
\end{aligned}$$

with $\mathfrak{Im}[\cdot]$ standing for the imaginary part of a complex number. Let us first analyze $\mathcal{B}_{N,\boldsymbol{W}}^{(1)}$. Using decomposition (3), we obtain that

$$\frac{1}{N}\sum_{n=0}^{N-1}\boldsymbol{r}_y(n)e^{-2i\pi\alpha_0 n} = \boldsymbol{r}_y^{(\alpha_0)} + \varepsilon_1,$$

$$\varepsilon_1 := \boldsymbol{r}_y^{(0)}\,\frac{1}{N}\sum_{n=0}^{N-1}e^{-2i\pi\alpha_0 n} + \boldsymbol{r}_y^{(-\alpha_0)}\,\frac{1}{N}\sum_{n=0}^{N-1}e^{-4i\pi\alpha_0 n}. \tag{20}$$

In the same way

$$\begin{aligned}
\frac{1}{N\sqrt{N}}\sum_{n=0}^{N-1}\boldsymbol{r}_y(n)ne^{-2i\pi\alpha_0 n} &= \frac{N+1}{2\sqrt{N}}\boldsymbol{r}_y^{(\alpha_0)} + \varepsilon_2,\\
\varepsilon_2 := \boldsymbol{r}_y^{(0)}\,\frac{1}{N\sqrt{N}}\sum_{n=0}^{N-1}ne^{-2i\pi\alpha_0 n} &+ \boldsymbol{r}_y^{(-\alpha_0)}\,\frac{1}{N\sqrt{N}}\sum_{n=0}^{N-1}ne^{-4i\pi\alpha_0 n}.
\end{aligned}$$

The asymptotic behavior of the terms ε_1 and ε_2 is characterized by the following lemma.

Lemma 4: For every strictly positive scalar δ and any positive integer K

$$\frac{1}{N^{(K+\delta)}}\sum_{n=0}^{N-1}n^K e^{2i\pi\alpha n} \longrightarrow 0, \qquad \text{as } N\to\infty, \text{ for } \alpha\neq 0.$$

Lemma 4 implies that

$$\varepsilon_1 \to 0, \;\; \varepsilon_2 \to 0, \;\; \text{and} \;\; \sqrt{N}\varepsilon_1 \to 0, \qquad \text{as } N\to\infty. \tag{21}$$

It is also easy to verify that

$$\begin{aligned}
\mathcal{B}_{N,\boldsymbol{W}}^{(1)} = -2i\pi\Bigg[\frac{(N+1)}{2\sqrt{N}}\Big(&\boldsymbol{r}_y^{(\alpha_0)^*}\boldsymbol{W}\boldsymbol{r}_y^{(\alpha_0)} - \boldsymbol{r}_y^{(\alpha_0)^*}\boldsymbol{W}\boldsymbol{r}_y^{(\alpha_0)}\\
&+\varepsilon_1^*\boldsymbol{W}\boldsymbol{r}_y^{(\alpha_0)} - \boldsymbol{r}_y^{(\alpha_0)^*}\boldsymbol{W}\varepsilon_1\Big) + \boldsymbol{r}_y^{(\alpha_0)^*}\boldsymbol{W}\varepsilon_2\\
&-\varepsilon_2^*\boldsymbol{W}\boldsymbol{r}_y^{(\alpha_0)} + \varepsilon_1^*\boldsymbol{W}\varepsilon_2 - \varepsilon_2^*\boldsymbol{W}\varepsilon_1\Bigg].
\end{aligned}$$

One can observe that the imaginary part operator ($\mathfrak{Im}$) allows to cancel out the nonconvergent term $(N+1)\boldsymbol{r}_y^{(\alpha_0)^*}\boldsymbol{W}\boldsymbol{r}_y^{(\alpha_0)}/2\sqrt{N}$ present in the expression of $\mathcal{B}_{N,\boldsymbol{W}}^{(1)}$. Therefore, $\mathcal{B}_{N,\boldsymbol{W}}^{(1)}$ can be expressed as

$$\begin{aligned}
\mathcal{B}_{N,\boldsymbol{W}}^{(1)} = -2i\pi\Bigg[&\frac{N+1}{2\sqrt{N}}\left(\varepsilon_1^*\boldsymbol{W}\boldsymbol{r}_y^{(\alpha_0)} - \boldsymbol{r}_y^{(\alpha_0)^*}\boldsymbol{W}\varepsilon_1\right)\\
&+\boldsymbol{r}_y^{(\alpha_0)^*}\boldsymbol{W}\varepsilon_2 - \varepsilon_2^*\boldsymbol{W}\boldsymbol{r}_y^{(\alpha_0)} + \varepsilon_1^*\boldsymbol{W}\varepsilon_2 - \varepsilon_2^*\boldsymbol{W}\varepsilon_1\Bigg].
\end{aligned}$$

Using the properties (21) of ε_1 and ε_2, it follows that $\mathcal{B}_{N,\boldsymbol{W}}^{(1)}$ converges to zero.

We note also that $\mathcal{B}_{N,\boldsymbol{W}}^{(2)}$ can be expressed as

$$\mathcal{B}_{N,\boldsymbol{W}}^{(2)} = -2i\pi\boldsymbol{R}(N)\mathcal{W}\boldsymbol{E}(N) \tag{22}$$

where

$$\begin{aligned}
\mathcal{W} &:= \mathrm{diag}(-\boldsymbol{W}, \boldsymbol{W}, \overline{\boldsymbol{W}}, -\overline{\boldsymbol{W}})\\
\boldsymbol{R}(N) &:= [\boldsymbol{R}_1^*(N), \boldsymbol{R}_0^*(N), \boldsymbol{R}_1^T(N), \boldsymbol{R}_0^T(N)]^T
\end{aligned}$$

then

$$\boldsymbol{E}(N) := [\boldsymbol{E}_0^T(N), \boldsymbol{E}_1^T(N), \boldsymbol{E}_0^*(N), \boldsymbol{E}_1^*(N)]^T$$

and

$$\begin{aligned}\boldsymbol{E}_0(N) &:= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \boldsymbol{e}_2(n) e^{-2i\pi\alpha_0 n}\\ \boldsymbol{R}_0(N) &:= \frac{1}{N} \sum_{n=0}^{N-1} \boldsymbol{r}_y(n) e^{-2i\pi\alpha_0 n}\\ \boldsymbol{E}_1(N) &:= \frac{1}{N\sqrt{N}} \sum_{n=0}^{N-1} \boldsymbol{e}_2(n) n e^{-2i\pi\alpha_0 n}\\ \boldsymbol{R}_1(N) &:= \frac{1}{N^2} \sum_{n=0}^{N-1} \boldsymbol{r}_y(n) n e^{-2i\pi\alpha_0 n}.\end{aligned}$$

Using (3), we obtain that $\boldsymbol{R}(N)$ depends only on terms of the form

$$\frac{1}{N^{(K+1)}} \sum_{n=0}^{N-1} n^K \exp(2i\pi\alpha n)$$

with $K = 0, 1$, and $\alpha = 0, -\alpha_0, -2\alpha_0$. As $\alpha_0 > 0$, it is easy to show that

$$\boldsymbol{R}(N) \rightarrow \left[\tfrac{1}{2}\boldsymbol{r}_y^{(\alpha_0)^*}, \boldsymbol{r}_y^{(\alpha_0)^*}, \tfrac{1}{2}\boldsymbol{r}_y^{(\alpha_0)^T}, \boldsymbol{r}_y^{(\alpha_0)^T}\right]^T. \tag{23}$$

We prove in Appendix II that $\boldsymbol{E}(N)$ converges in distribution to a zero-mean Gaussian distribution. This result shows that $\mathcal{B}_{N,\boldsymbol{W}}^{(2)}$ is asymptotically zero-mean Gaussian. As for $\mathcal{B}_{N,\boldsymbol{W}}^{(3)}$, it can be expressed in the following form:

$$\mathcal{B}_{N,\boldsymbol{W}}^{(3)} = -4\pi \mathfrak{Im}\left[\boldsymbol{E}_1^*(N)\boldsymbol{W}\left(\frac{1}{N}\sum_{n=0}^{N-1} \boldsymbol{e}_2(n) e^{-2i\pi\alpha_0 n}\right)\right].$$

Lemma 1 shows that

$$\frac{1}{N}\sum_{n=0}^{N-1} \boldsymbol{e}_2(n) \exp(2i\pi\alpha_0 n)$$

converges almost surely to zero. Since $\boldsymbol{E}_1(N)$ is a component of the vector $\boldsymbol{E}(N)$, it is asymptotically Gaussian. This implies that $\mathcal{B}_{N,\boldsymbol{W}}^{(3)}$ almost surely converges to zero [5].

According to (19), we deduce the following result.

Proposition 2:

$$\mathcal{B}_{N,\boldsymbol{W}} \rightarrow \mathcal{N}(0, \gamma_{\mathcal{B}}) \text{ in distribution,} \qquad \text{as } N \rightarrow \infty$$

where

$$\gamma_{\mathcal{B}} := \lim_{N\rightarrow\infty} \mathbb{E}\left[\mathcal{B}_{N,\boldsymbol{W}}^{(2)} \mathcal{B}_{N,\boldsymbol{W}}^{(2)^*}\right].$$

Using Propositions 1 and 2 as well as (18), we finally obtain the main result of this section.

Theorem 2:

$$N^{\frac{3}{2}}(\hat{\alpha}_{N,\boldsymbol{W}} - \alpha_0) \rightarrow \mathcal{N}(0, \gamma_{\boldsymbol{W}}) \text{ in distribution, as } N \rightarrow \infty$$

where $\gamma_{\boldsymbol{W}} := \gamma_{\mathcal{A}}^{-1}\gamma_{\mathcal{B}}\gamma_{\mathcal{A}}^{-1}$.

We have thus proved the asymptotic normality of $N^{\frac{3}{2}}\delta\hat{\alpha}_{N,\boldsymbol{W}}$, and that the convergence rate of $\delta\hat{\alpha}_{N,\boldsymbol{W}}$ is $N^{-3/2}$ as in standard frequency estimation problems.

3) Computation of $\gamma_{\boldsymbol{W}}$: We now compute in closed form $\gamma_{\boldsymbol{W}} := \gamma_{\mathcal{A}}^{-1}\gamma_{\mathcal{B}}\gamma_{\mathcal{A}}^{-1}$. As $\gamma_{\mathcal{A}}$ has been evaluated (see Proposition 1), it remains to calculate a closed-form expression for

$$\gamma_{\mathcal{B}} := \lim_{N\rightarrow\infty} \mathbb{E}\left[\mathcal{B}_{N,\boldsymbol{W}}^{(2)} \mathcal{B}_{N,\boldsymbol{W}}^{(2)^*}\right].$$

Using (22) and (23), it is easy to check that

$$\begin{aligned}\gamma_{\mathcal{B}} = 4\pi^2 &\left[\frac{1}{2}\boldsymbol{r}_y^{(\alpha_0)^*}, \boldsymbol{r}_y^{(\alpha_0)^*}, \frac{1}{2}\boldsymbol{r}_y^{(\alpha_0)^T}, \boldsymbol{r}_y^{(\alpha_0)^T}\right]\mathcal{W}\\ &\times \left(\lim_{N\rightarrow\infty}\mathbb{E}[\boldsymbol{E}(N)\boldsymbol{E}^*(N)]\right)\\ &\times \mathcal{W}\left[\frac{1}{2}\boldsymbol{r}_y^{(\alpha_0)^*}, \boldsymbol{r}_y^{(\alpha_0)^*}, \frac{1}{2}\boldsymbol{r}_y^{(\alpha_0)^T}, \boldsymbol{r}_y^{(\alpha_0)^T}\right]^*.\end{aligned}$$

In order to evaluate $\lim_{N\rightarrow\infty}\mathbb{E}[\boldsymbol{E}(N)\boldsymbol{E}^*(N)]$, we observe that

$$\begin{aligned}&\mathbb{E}[\boldsymbol{E}(N)\boldsymbol{E}^*(N)]\\ &= \begin{bmatrix} \boldsymbol{P}_N(0,0) & \boldsymbol{P}_N(0,1) & \boldsymbol{P}_N^{(c)}(0,0) & \boldsymbol{P}_N^{(c)}(0,1)\\ \boldsymbol{P}_N(1,0) & \boldsymbol{P}_N(1,1) & \boldsymbol{P}_N^{(c)}(1,0) & \boldsymbol{P}_N^{(c)}(1,1)\\ \overline{\boldsymbol{P}}_N^{(c)}(0,0) & \overline{\boldsymbol{P}}_N^{(c)}(0,1) & \overline{\boldsymbol{P}}_N(0,0) & \overline{\boldsymbol{P}}_N(0,1)\\ \overline{\boldsymbol{P}}_N^{(c)}(1,0) & \overline{\boldsymbol{P}}_N^{(c)}(1,1) & \overline{\boldsymbol{P}}_N(1,0) & \overline{\boldsymbol{P}}_N(1,1)\end{bmatrix}\end{aligned} \tag{24}$$

where

$$\begin{aligned}\boldsymbol{P}_N(K, K') &:= \mathbb{E}[\boldsymbol{E}_K(N)\boldsymbol{E}_{K'}^*(N)]\\ &= \frac{1}{N^{(K+K'+1)}}\sum_{\substack{n=0\\n'=0}}^{N-1}\mathbb{E}[\boldsymbol{e}_2(n)\boldsymbol{e}_2^*(n')]n^K n'^{K'}\\ &\quad\times e^{-2i\pi\alpha_0 n}e^{2i\pi\alpha_0 n'}\end{aligned}$$

and:

$$\begin{aligned}\boldsymbol{P}_N^{(c)}(K, K') &:= \mathbb{E}[\boldsymbol{E}_K(N)\boldsymbol{E}_{K'}^T(N)]\\ &= \frac{1}{N^{(K+K'+1)}}\sum_{\substack{n=0\\n'=0}}^{N-1}\mathbb{E}[\boldsymbol{e}_2(n)\boldsymbol{e}_2^T(n')]n^K n'^{K'}\\ &\quad\times e^{-2i\pi\alpha_0 n}e^{-2i\pi\alpha_0 n'}.\end{aligned}$$

We now study the asymptotic behavior of these terms. For this, we have to specify the properties of $\boldsymbol{e}_2(n)$. It is cyclostationary, and using some results of [15], [17], and [22], we obtain that its set of significant cyclic frequencies is given[2] by $\mathcal{F} = \{l\alpha_0 \bmod 1 \mid |l| \leq 3\}$. In other words

$$\boldsymbol{R}_{\boldsymbol{e}_2}(n, \tau) = \mathbb{E}[\boldsymbol{e}_2(n+\tau)\boldsymbol{e}_2^*(n)]$$

and

$$\boldsymbol{R}_{\boldsymbol{e}_2}^{(c)}(n, \tau) = \mathbb{E}[\boldsymbol{e}_2(n+\tau)\boldsymbol{e}_2(n)]$$

can be expressed as

$$\begin{aligned}\boldsymbol{R}_{\boldsymbol{e}_2}(n, \tau) &= \sum_{\alpha\in\mathcal{F}} \boldsymbol{r}_e^{(\alpha)} e^{2i\pi n\alpha},\\ \boldsymbol{R}_{\boldsymbol{e}_2}^{(c)}(n, \tau) &= \sum_{\alpha\in\mathcal{F}} \boldsymbol{r}_e^{c,(\alpha)} e^{2i\pi n\alpha}.\end{aligned}$$

[2]The notation $x \bmod 1$ stands for the value of x modulo 1.

Using Assumption 1 and well-known results on Césaro sums (see, e.g., [19]), we obtain after some simple manipulations that

$$\boldsymbol{P}_N(K, K') \to \frac{1}{K+K'+1} S_{\boldsymbol{e}_2}^{(0)}(e^{2i\pi\alpha_0}), \qquad \text{as } N \to \infty \tag{25}$$

$$\boldsymbol{P}_N^{(c)}(K, K') \to \frac{1}{K+K'+1} S_{\boldsymbol{e}_2^{(c)}}^{(2\alpha_0)}(e^{2i\pi\alpha_0}), \qquad \text{as } N \to \infty \tag{26}$$

where $S_{\boldsymbol{e}_2}^{(0)}(\exp(2i\pi f))$ and $S_{\boldsymbol{e}_2^{(c)}}^{(2\alpha_0)}(\exp(2i\pi f))$ represent the cyclic spectrum at cyclic frequency 0 and the conjugate cyclic spectrum at cyclic frequency $2\alpha_0$ of $\boldsymbol{e}_2(n)$, respectively. Fortunately, the spectra $S_{\boldsymbol{e}_2}^{(0)}(\exp(2i\pi\alpha_0))$ and $S_{\boldsymbol{e}_2^{(c)}}^{(2\alpha_0)}(\exp(2i\pi\alpha_0))$ can be expressed more explicitly. Let

$$\delta\hat{\boldsymbol{r}}_N^{(\alpha_0)} := \hat{\boldsymbol{r}}_N^{(\alpha_0)} - \boldsymbol{r}_y^{(\alpha_0)}$$

denote the estimation error corresponding to $\boldsymbol{r}_y^{(\alpha_0)}$. It is well known that $N^{1/2}(\delta\hat{\boldsymbol{r}}_N^{(\alpha_0)})$ converges to a Gaussian distribution [7]. Let $\boldsymbol{\Gamma}$ and $\boldsymbol{\Gamma}^{(c)}$ denote the unconjugated/conjugated asymptotic covariance matrices

$$\boldsymbol{\Gamma} := \lim_{N\to\infty} N\mathbb{E}\left[\delta\hat{\boldsymbol{r}}_N^{(\alpha_0)}\delta\hat{\boldsymbol{r}}_N^{(\alpha_0)^*}\right]$$
$$\boldsymbol{\Gamma}^{(c)} := \lim_{N\to\infty} N\mathbb{E}\left[\delta\hat{\boldsymbol{r}}_N^{(\alpha_0)}\delta\hat{\boldsymbol{r}}_N^{(\alpha_0)^T}\right].$$

In order to express $\boldsymbol{\Gamma}$ and $\boldsymbol{\Gamma}^{(c)}$ in terms of $S_{\boldsymbol{e}_2}^{(0)}(\exp(2i\pi\alpha_0))$ and $S_{\boldsymbol{e}_2^{(c)}}^{(2\alpha_0)}(\exp(2i\pi\alpha_0))$, we note that since

$$\boldsymbol{y}_2(n) = \sum_{k=-1}^{1} \boldsymbol{r}_y^{(k\alpha_0)} \exp(2i\pi k\alpha_0 n) + \boldsymbol{e}_2(n)$$

it follows that

$$\hat{\boldsymbol{r}}_N^{(\alpha_0)} = \boldsymbol{R}_0(N) + \frac{1}{\sqrt{N}}\boldsymbol{E}_0(N).$$

Substituting ε_1 from (20), we obtain that

$$\sqrt{N}\,\delta\hat{\boldsymbol{r}}_N^{(\alpha_0)} = \boldsymbol{E}_0(N) + \sqrt{N}\,\varepsilon_1.$$

From (21), it follows that $\sqrt{N}\varepsilon_1$ is a deterministic term converging to zero, as $N \to \infty$. From this, we deduce further that [5]

$$\begin{aligned}\boldsymbol{\Gamma} &= \lim_{N\to\infty} \mathbb{E}[\boldsymbol{E}_0(N)\boldsymbol{E}_0^*(N)] = \lim_{N\to\infty} \boldsymbol{P}_N(0, 0)\\ &= S_{\boldsymbol{e}_2}^{(0)}(e^{2i\pi\alpha_0})\end{aligned} \tag{27}$$

$$\begin{aligned}\boldsymbol{\Gamma}^{(c)} &= \lim_{N\to\infty} \mathbb{E}[\boldsymbol{E}_0(N)\boldsymbol{E}_0^T(N)] = \lim_{N\to\infty} \boldsymbol{P}_N^{(c)}(0, 0)\\ &= S_{\boldsymbol{e}_2^{(c)}}^{(2a_0)}(e^{2i\pi\alpha_0}).\end{aligned} \tag{28}$$

Plugging (27), (28), (25), and (26) back into (24) yields

$$\lim_{N\to\infty} \mathbb{E}[\boldsymbol{E}(N)\boldsymbol{E}^*(N)] = \begin{bmatrix} \boldsymbol{\Gamma} & \frac{1}{2}\boldsymbol{\Gamma} & \boldsymbol{\Gamma}^{(c)} & \frac{1}{2}\boldsymbol{\Gamma}^{(c)} \\ \frac{1}{2}\boldsymbol{\Gamma} & \frac{1}{3}\boldsymbol{\Gamma} & \frac{1}{2}\boldsymbol{\Gamma}^{(c)} & \frac{1}{3}\boldsymbol{\Gamma}^{(c)} \\ \overline{\boldsymbol{\Gamma}}^{(c)} & \frac{1}{2}\overline{\boldsymbol{\Gamma}}^{(c)} & \overline{\boldsymbol{\Gamma}} & \frac{1}{2}\overline{\boldsymbol{\Gamma}} \\ \frac{1}{2}\overline{\boldsymbol{\Gamma}}^{(c)} & \frac{1}{3}\overline{\boldsymbol{\Gamma}}^{(c)} & \frac{1}{2}\overline{\boldsymbol{\Gamma}} & \frac{1}{3}\overline{\boldsymbol{\Gamma}} \end{bmatrix}. \tag{29}$$

Finally, by combining all the previous results, we obtain the following compact and simple expression for the asymptotic variance:

$$\begin{aligned}\gamma_{\boldsymbol{W}} = \frac{3}{\pi^2}\left(\boldsymbol{R}^{(\alpha_0)^*}\tilde{\boldsymbol{W}}^*\boldsymbol{R}^{(\alpha_0)}\right)^{-1}\boldsymbol{R}^{(\alpha_0)^*}\tilde{\boldsymbol{W}}\boldsymbol{G}\tilde{\boldsymbol{W}}\boldsymbol{R}^{(\alpha_0)}\\ \cdot\left(\boldsymbol{R}^{(\alpha_0)^*}\tilde{\boldsymbol{W}}\boldsymbol{R}^{(\alpha_0)}\right)^{-1}\end{aligned} \tag{30}$$

where

$$\boldsymbol{R}^{(\alpha_0)} := \begin{bmatrix}\boldsymbol{r}_y^{(\alpha_0)}\\ \overline{\boldsymbol{r}}_y^{(\alpha_0)}\end{bmatrix}, \quad \tilde{\boldsymbol{W}} := \begin{bmatrix}\boldsymbol{W} & 0\\ 0 & \overline{\boldsymbol{W}}\end{bmatrix}$$
$$\boldsymbol{G} := \begin{bmatrix}\boldsymbol{\Gamma} & -\boldsymbol{\Gamma}^{(c)}\\ -\overline{\boldsymbol{\Gamma}}^{(c)} & \overline{\boldsymbol{\Gamma}}\end{bmatrix}.$$

IV. Choosing the Weighting Matrix $\boldsymbol{W}$ and Parameter Υ

We now exploit (30) to study the influence of $\boldsymbol{W}$ and Υ on the performance of the estimate. A natural question that arises is: how do we choose the weighting matrix $\boldsymbol{W}$ in order to improve the estimation? If $\boldsymbol{G}$ is invertible, we can observe (see, e.g., [19]) that for any $\boldsymbol{W} = \boldsymbol{W}^* \geq 0$, the left-hand side (LHS) of (30) is less than $(3/\pi^2)(\boldsymbol{R}^{(\alpha_0)^*}\boldsymbol{G}^{-1}\boldsymbol{R}^{(\alpha_0)})^{-1}$, which represents the value of the right-hand side (RHS) of (30) evaluated for $\tilde{\boldsymbol{W}} = \boldsymbol{G}^{-1}$. This result can be extended even when $\boldsymbol{G}$ is a singular matrix by replacing the regular inverse $\boldsymbol{G}^{-1}$ with the (Moore–Penrose) pseudo-inverse $\boldsymbol{G}^{\#}$ (see, e.g., [1]). However, the matrix $\boldsymbol{G}^{\#}$ is not a block-diagonal matrix in general, i.e., it may happen that there is no weighting matrix $\boldsymbol{W}$ for which $\gamma_{\boldsymbol{W}}$ coincides with $(3/\pi^2)(\boldsymbol{R}^{(\alpha_0)^*}\boldsymbol{G}^{\#}\boldsymbol{R}^{(\alpha_0)})^{-1}$. Therefore, the matrix inequality trick used in [19] to derive optimal weighting matrices is not applicable here. Hence, determining a positive-definite optimal weighting matrix seems to be a difficult problem.

We show next that $\boldsymbol{W} = \boldsymbol{I}$ seems to be an appropriate choice, at least for high singgal-to-noise ratios (SNRs). Indeed, we prove in the sequel that $\boldsymbol{W} = \boldsymbol{I}$ leads to a very low variance estimate if Υ is chosen large enough. For this, we assume for the sake of simplicity that the sequence s_n is circular: in this case, the expressions of $\boldsymbol{\Gamma}$ and $\boldsymbol{\Gamma}^{(c)}$ are simple. Using that the spectrum of $y_a(t)$ is limited to the interval $[-(1+\rho)/(2T_s), (1+\rho)/(2T_s)]$ for some parameter $0<\rho<1$ and $T_e<T_s/4$, one can show that (see Appendix III)

$$\begin{aligned}\boldsymbol{\Gamma} = \int_0^1 S_y^{(0)}(e^{2i\pi f})\overline{S}_y^{(0)}\left(e^{2i\pi(f-\alpha_0)}\right)\boldsymbol{d}_\Upsilon(e^{2i\pi f})\\ \times \boldsymbol{d}_\Upsilon^*(e^{2i\pi f})\,df + \frac{\kappa}{\alpha_0}\boldsymbol{r}_y^{(\alpha_0)}\boldsymbol{r}_y^{(\alpha_0)^*}\end{aligned} \tag{31}$$

$$\begin{aligned}\boldsymbol{\Gamma}^{(c)} = \int_0^1 \left(S_y^{(\alpha_0)}(e^{2i\pi f})\right)^2\boldsymbol{d}_\Upsilon(e^{2i\pi f})\boldsymbol{d}_\Upsilon^T(e^{2i\pi f})\,df\\ + \frac{\kappa}{\alpha_0}\boldsymbol{r}_y^{(\alpha_0)}\boldsymbol{r}_y^{(\alpha_0)^T}\end{aligned} \tag{32}$$

where

$$\boldsymbol{d}_\Upsilon(e^{2i\pi f}) := [\exp(-2i\pi\Upsilon f), \ldots, \exp(2i\pi\Upsilon f)]^T$$

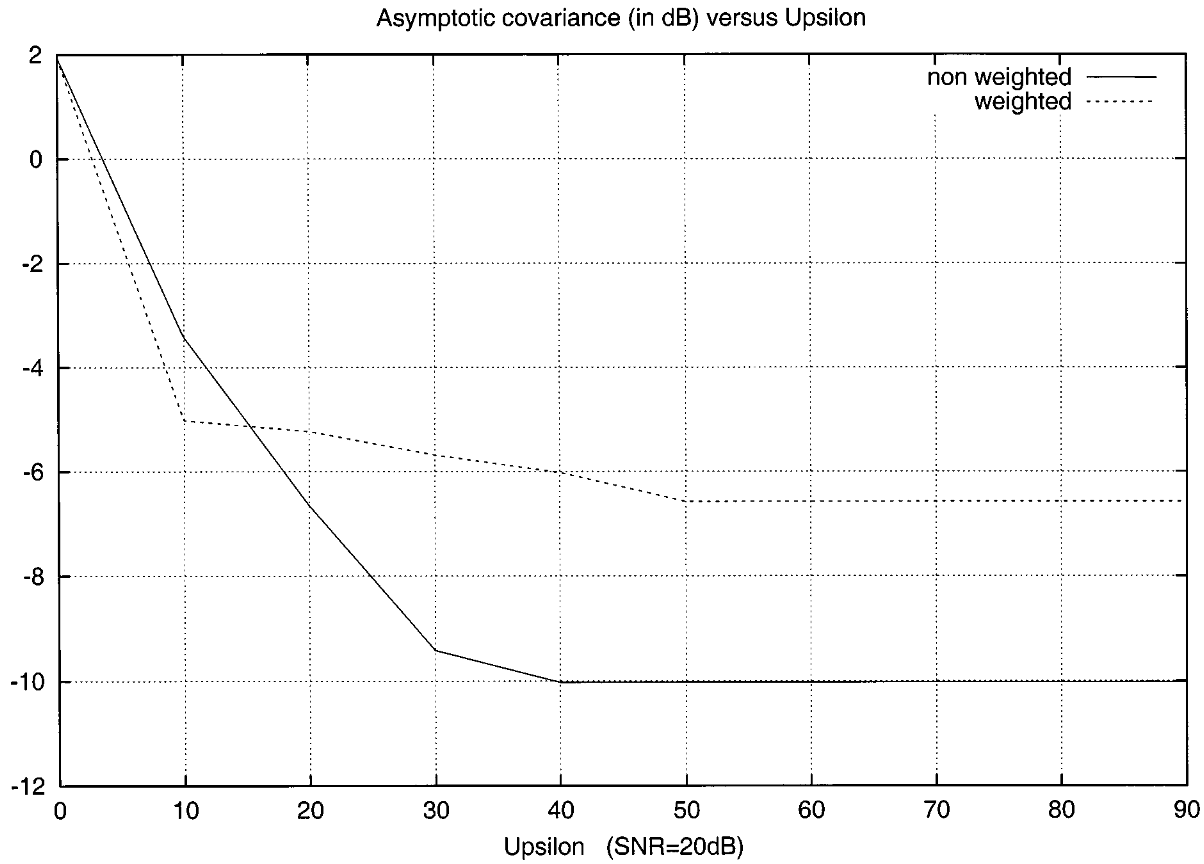

Fig. 1. $\gamma_{\boldsymbol{I}}$ and $\gamma_{\boldsymbol{\Gamma}^{\#}}$ (in decibels) versus Υ.

and κ is the kurtosis of $\{s_n\}_{n\in\mathbb{Z}}$. Using the condition $T_e < T_s/4$, it is easy to show that

$$S_y^{(0)}(e^{2i\pi f}) = \frac{T_e}{T_s}\left|h(e^{2i\pi f})\right|^2 + S_w(e^{2i\pi f})$$
$$S_y^{(\alpha_0)}(e^{2i\pi f}) = \frac{T_e}{T_s}h(e^{2i\pi f})h^*\left(e^{2i\pi(f-\alpha_0)}\right) \qquad (33)$$

where $h(z) := \sum_k h_a(kT_e)z^{-k}$ and $S_w(\exp(2i\pi f))$ stands for the spectral density of $w(n) := w_a(nT_e)$.

Let us now consider the noiseless case. The product $S_y^{(0)}(\exp(2i\pi f))\ \overline{S}_y^{(0)}(\exp(2i\pi(f-\alpha_0)))$ coincides with $|S_y^{(\alpha_0)}(\exp(2i\pi f))|^2$. After some straightforward manipulations, we obtain that the asymptotic variance of the estimate $\hat{\alpha}_{N,\boldsymbol{I}}$ is given by

$$\gamma_{\boldsymbol{I}} = \frac{3\Phi(\Upsilon)}{4\pi^2\|\boldsymbol{r}_y^{(\alpha_0)}\|^4} \qquad (34)$$

where

$$\Phi(\Upsilon) = 2\int_0^1 \left|S_y^{(\alpha_0)}(e^{2i\pi f})\right|^2 \left|\sum_{\tau=-\Upsilon}^{\Upsilon} r_y^{(\alpha_0)}(\tau)e^{-2i\pi f\tau}\right|^2 df$$
$$- 2\mathfrak{Re}\left\{\int_0^1 \left[\overline{S}_y^{(\alpha_0)}(e^{2i\pi f})\right]^2 \times \left[\sum_{\tau=-\Upsilon}^{\Upsilon} r_y^{(\alpha_0)}(\tau)e^{-2i\pi f\tau}\right]^2 df\right\}$$

where $\mathfrak{Re}[\cdot]$ stands for the real part of a complex number. The important point is to observe that if Υ is large enough

$$\sum_{\tau=-\Upsilon}^{\Upsilon} r_y^{(\alpha_0)}(\tau)e^{-2i\pi f\tau} \simeq S_y^{(\alpha_0)}(e^{2i\pi f})$$

and that $\Phi(\Upsilon) \simeq 0$. It turns out that, in the noiseless case, the asymptotic variance of $\hat{\alpha}_{N,\boldsymbol{I}}$ converges to 0 as Υ increases. We note that this result holds because we have chosen $\boldsymbol{W} = \boldsymbol{I}$. In particular, it seems that if $\boldsymbol{W} \neq \boldsymbol{I}$, the estimate $\hat{\alpha}_{N,\boldsymbol{W}}$ does not have the above property. Therefore, the choice $\boldsymbol{W} = \boldsymbol{I}$ seems reasonable, at least if the SNR is high enough and Υ is chosen large enough.

Remark: In practice, the maximization of $J_{N,\boldsymbol{W}}(\alpha)$ eith respect to (w.r.t.) α is not so easy to achieve because it shows several spurious local maxima. The estimate $\hat{\alpha}_{N,\boldsymbol{W}}$ is thus computed in two steps. In the first step (coarse search), $J_{N,\boldsymbol{W}}(\alpha)$ is evaluated by means of a fast Fourier transform (FFT) algorithm on the grid $\{a_k = k/N,\ k = 0, \ldots, N-1\}$. In the second step (fine search), a gradient maximization algorithm of $J_{N,\boldsymbol{W}}(\alpha)$, initialized at the argument of $\max_{a_k} J_{N,\boldsymbol{W}}(a_k)$, gives the estimate $\hat{\alpha}_{N,\boldsymbol{W}}$. As the asymptotic analysis presented in this paper is closely related to the local behavior of $J_{N,\boldsymbol{W}}(\alpha)$ around α_0, it of course only allows to predicate the performance of the fine search, provided the coarse search has been successful. In particular, that $\boldsymbol{W} = I$ is relevant in the context of the fine search does not imply that it is well adapted to the coarse search. Indeed, let $\boldsymbol{\Gamma}(\alpha)$ be the asymptotic covariance matrix of $\delta\hat{\boldsymbol{r}}_N^{(\alpha)}$, so

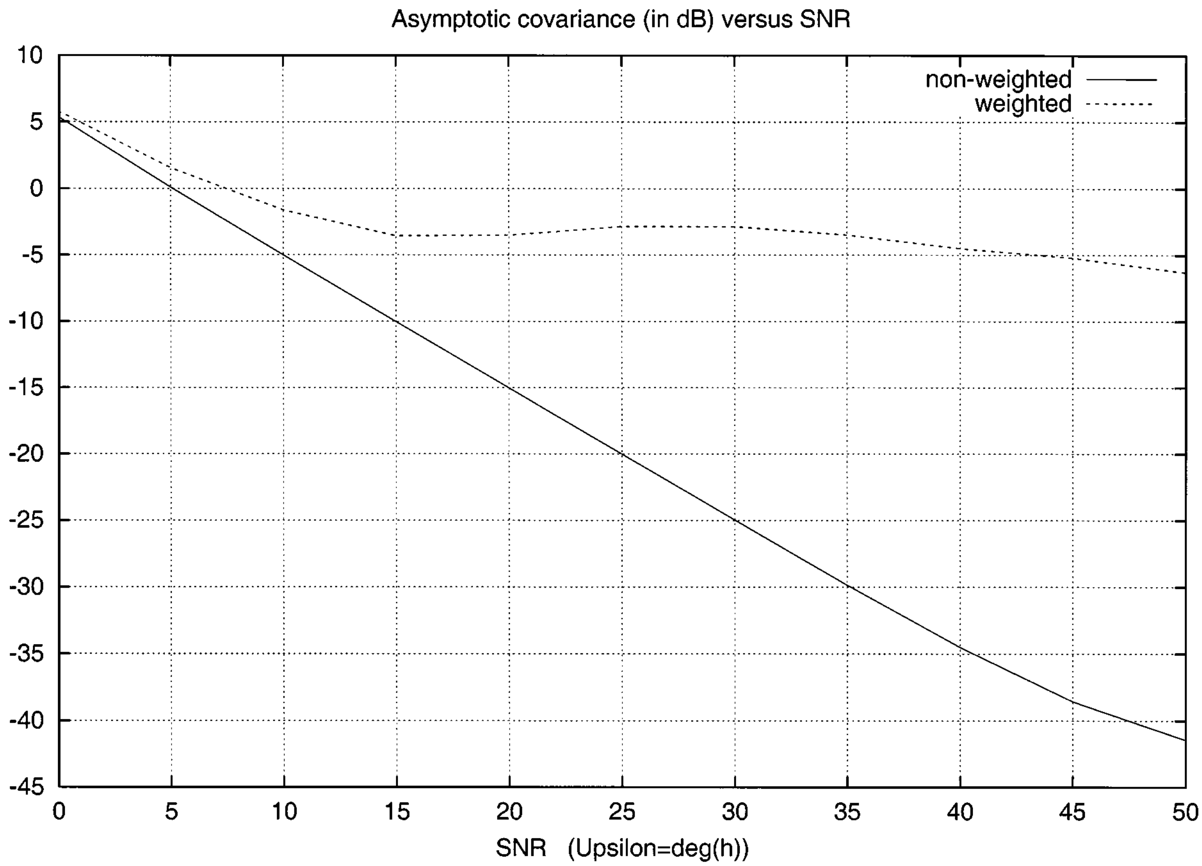

Fig. 2. $\gamma_{\boldsymbol{I}}$ and $\gamma_{\boldsymbol{\Gamma}^{\#}}$ (in decibels) versus SNR, for $\Upsilon = M$.

that $\Gamma = \Gamma(\alpha_0)$. It has been shown in [6] and [18] that choosing $\boldsymbol{W} = \Gamma(\alpha)^{\#}$ allows, in principle, to improve dramatically the performance of the coarse search. However, in the context of the fine search, weighting the cost function $J_{N,\boldsymbol{W}}(\alpha)$ by $\boldsymbol{W} = \boldsymbol{\Gamma}^{\#}$ is not necessarily recommended. This is going to be confirmed by the numerical evaluations in the following section.

V. Numerical Illustrations

We now illustrate the above results by some numerical evaluations of the asymptotic variance of $\hat{\alpha}_{N,\boldsymbol{W}}$. We assume that the shaping filter used by the transmitter is a square-root raised-cosine pulse shape with excess bandwidth $\rho = 0.2$, and that the propagation channel is a multipath channel. The amplitude, phases, and time delays of the channel are random variables, and each curve is obtained by averaging the variances of our estimate over 100 different realizations of the propagation channel. It is assumed that $\alpha_0 = T_e/T_s$ is equal to $1/5$. In order to obtain quasi-band-limited signals, we have used a degree $M = 80$ polynomial $h(z)$.

In Fig. 1, the noise $w(n)$ is white, SNR is equal to 20 dB, and we study the influence of Υ on the asymptotic covariance $\gamma_{\boldsymbol{W}}$.

We compare the unweighted estimate ($\boldsymbol{W} = \boldsymbol{I}$) and the weighted estimate corresponding to the choice $\boldsymbol{W} = \boldsymbol{\Gamma}^{\#}$. We notice that the number of lags Υ taken into account has a great influence on the variance.

We now study the behavior of the asymptotic covariance of the estimate versus SNR. In Fig. 2, for sake of clarity, we only plot the case $\Upsilon = M$.

The variance of the weighted estimator does not converge to 0 when SNR increases, while the curve of the unweighted estimator confirms our calculations. Indeed, the variance of $\hat{\alpha}_{N,\boldsymbol{I}}$ decreases to 0 as Υ increases. Moreover, the unweighted estimator has a lower variance for Υ large enough.

VI. Conclusion

In this paper, we have studied rigorously the asymptotic performance of a symbol rate estimator. We have shown that the estimator is consistent, asymptotically normal, and that its convergence rate is $N^{-3/2}$. Our approach also leads to an interpretable closed-form expression for the asymptotic variance. We have taken advantage of our interpretable formula to discuss guidelines in selecting certain important parameters.

Appendix I
Proof of Lemma 1

We first note that $\boldsymbol{s}_N^{(K)}(\alpha)$ satisfies the statement of Lemma 1 if and only if any of its components verifies (10). In this appendix, for brevity, we only prove the convergence of

$$[\boldsymbol{s}_N^{(K)}(\alpha)]_0 = \frac{1}{N^{(K+1)}} \sum_{n=0}^{N-1} n^K \boldsymbol{e}_0^2(n) \exp(2i\pi\alpha n).$$

The reader may check that the other components can be treated using similar arguments. In the sequel, we denote $e(n) := \boldsymbol{e}_0(n)$

and $s_N^{(K)}(\alpha) := [\boldsymbol{s}_N^{(K)}(\alpha)]_0$. We will rely on Hannan's proof [13] to show that

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

We first study the second-order moment of $\sup_{\alpha\in[0,1]}|s_N^{(K)}(\alpha)|$ which is given by

$$\begin{aligned}V_1 &:= \mathbb{E}\left[\sup_{\alpha\in[0,1]}\left|s_N^{(K)}(\alpha)\right|^2\right]\\ &= \mathbb{E}\left[\sup_{\alpha\in[0,1]}\frac{1}{N^{2(K+1)}}\sum_{n,\,n'=0}^{N-1} n^K n'^K e(n)e^*(n')\right.\\ &\qquad\left.\times e^{2i\pi(n-n')\alpha}\right].\end{aligned}$$

Defining $A_n := \sup(-n;\,0)$ and $B_n := \inf(N-1-n;\,N-1)$, it follows that

$$\begin{aligned}V_1 = \mathbb{E}&\left[\sup_{\alpha\in[0,1]}\frac{1}{N^{2(K+1)}}\sum_{n=-N+1}^{N-1} e^{2i\pi n\alpha}\sum_{n'=A_n}^{B_n}\right.\\ &\qquad\left.\times(n+n')^K n'^K e(n+n')e^*(n')\right].\end{aligned}$$

By using the triangular inequality, we obtain that

$$\begin{aligned}V_1 \le &\frac{1}{N^{2(K+1)}}\sum_{n=-N+1}^{N-1}\\ &\times\mathbb{E}\left[\left|\sum_{n'=A_n}^{B_n}(n+n')^K n'^K e(n+n')e^*(n')\right|\right].\end{aligned}$$

According to Schwartz inequality, it follows also that

$$\begin{aligned}&\mathbb{E}\left[\left|\sum_{n'=A_n}^{B_n}(n+n')^K n'^K e(n+n')e^*(n')\right|\right]\\ &\le\left(\mathbb{E}\left[\left|\sum_{n'=A_n}^{B_n}(n+n')^K n'^K e(n+n')e^*(n')\right|^2\right]\right)^{\frac12}\\ &=\left(\sum_{n_1,\,n_2=A_n}^{B_n}(n+n_1)^K.n_1{}^K.(n+n_2)^K.n_2{}^K\right.\\ &\qquad\left.\times\mathbb{E}\left[e(n+n_1)e^*(n_1)e^*(n+n_2)e(n_2)\right]\right)^{\frac12}.\end{aligned}$$

By taking advantage of the concavity of the square-root function and Jensen's inequality, we obtain that

$$\begin{aligned}V_1 \le &\frac{(2N+1)^{\frac12}}{N^{2(K+1)}}\\ &\times\left(\sum_{n=-N+1}^{N-1}\sum_{\substack{n_1=A_n\\ n_2=A_n}}^{B_n}(n+n_1)^K n_1{}^K(n+n_2)^K n_2{}^K\right.\\ &\qquad\left.\times\mathbb{E}\left[e(n+n_1)e^*(n_1)e^*(n+n_2)e(n_2)\right]\right)^{\frac12}.\end{aligned}$$

Therefore,

$$\begin{aligned}V_1 &= \mathbb{E}\left[\sup_{\alpha\in[0,1]}\left|s_N^{(K)}(\alpha)\right|^2\right]\\ &\le\frac{(2N+1)^{\frac12}}{N^{2(K+1)}}\left(a(N)+b(N)+c(N)+d(N)\right)^{\frac12}\end{aligned}\tag{35}$$

with

$$\begin{aligned}a(N) &:= \sum_{n=-N+1}^{N-1}\sum_{\substack{n_1=A_n\\ n_2=A_n}}^{B_n}(n+n_1)^K n_1{}^K(n+n_2)^K n_2{}^K\\ &\qquad\times r_e(n_1,\,n)r_e^*(n_2,\,n)\\ b(N) &:= \sum_{n=-N+1}^{N-1}\sum_{\substack{n_1=A_n\\ n_2=A_n}}^{B_n}(n+n_1)^K n_1{}^K(n+n_2)^K n_2{}^K\\ &\qquad\times r_e(n+n_2,\,n_1-n_2)r_e^*(n_2,\,n_1-n_2)\\ c(N) &:= \sum_{n=-N+1}^{N-1}\sum_{\substack{n_1=A_n\\ n_2=A_n}}^{B_n}(n+n_1)^K n_1{}^K(n+n_2)^K n_2{}^K\\ &\qquad\times r_e^c(n_2,\,n+n_1-n_2)r_e^{c*}(n_1,\,n+n_2-n_1)\\ d(N) &:= \sum_{n=-N+1}^{N-1}\sum_{\substack{n_1=A_n\\ n_2=A_n}}^{B_n}(n+n_1)^K n_1{}^K(n+n_2)^K n_2{}^K\\ &\qquad\times \mathrm{cum}(e(n+n_1),\,e^*(n_1),\,e^*(n+n_2),\,e(n_2))\end{aligned}$$

where

$$r_e(n,\,\tau) := \mathbb{E}[e(n+\tau)e^*(n)]$$

and

$$r_e^c(n,\,\tau) := \mathbb{E}[e(n+\tau)e(n)].$$

We obtain easily that

$$|a(N)| \le 4^K N^{4K}\sum_{n=-N+1}^{N-1}\sum_{n_1,\,n_2=0}^{N-1}|r_e(n_1,\,n)|\,|r_e(n_2,\,n)|.$$

From Assumption 1, we deduce that

$$|a(N)| \le 4^K N^{4K}\mathcal{M}_2^2(2N).$$

Thus,

$$a(N) = \mathcal{O}(N^{2(2K+1/2)}).$$

In the same way, we obtain that

$$b(N) = \mathcal{O}(N^{2(2K+1)})$$
$$c(N) = \mathcal{O}(N^{2(2K+1)})$$

and

$$d(N) = \mathcal{O}(N^{2(2K+1/2)}).$$

According to (35), it follows that

$$\mathbb{E}\left[\sup_{\alpha\in[0,1]}\left|s_N^{(K)}(\alpha)\right|^2\right] = \mathcal{O}\left(\frac{1}{N^{\frac12}}\right).$$

Let us consider $N(M)$, the smallest integer greater than $M^{2(1+\delta)}$ with $\delta > 0$. It follows that

$$\mathbb{E}\left[\sup_{\alpha\in[0,1]}\left|s_{N(M)}^{(K)}(\alpha)\right|^2\right] = \mathcal{O}\left(\frac{1}{M^{1+\delta}}\right).$$

Chebychev's inequality leads further to

$$\mathrm{Prob}\left(\sup_{\alpha\in[0,1]}\left|s_{N(M)}^{(K)}(\alpha)\right| > \varepsilon\right) \le \frac{1}{\varepsilon^2}\,\mathcal{O}\left(\frac{1}{M^{1+\delta}}\right).$$

Therefore, for any $\varepsilon > 0$

$$\sum_{M\in\mathbb{N}} \text{Prob}\left(\sup_{\alpha\in[0,1]} \left|s_{N(M)}^{(K)}(\alpha)\right| > \varepsilon\right) < +\infty.$$

The Borel–Cantelli lemma thus implies that

$$\text{Prob}\left(\lim_{M\to\infty} \sup_{\alpha\in[0,1]} \left|s_{N(M)}^{(K)}(\alpha)\right| = 0\right) = 1.$$

Thus,

$$\sup_{\alpha\in[0,1]} |s_{N(M)}^{(K)}(\alpha)| \xrightarrow{\text{a.s.}} 0, \qquad \text{as } M\to\infty.$$

We have only proved the result for a certain subsequence extracted from $s_N^{(K)}(\alpha)$. It remains to prove the uniform convergence of the sequence itself. For this, let N be an integer between $N(M)$ and $N(M+1)$ and consider the expression

$$\sup_{N(M)<N\le N(M+1)} \sup_{\alpha\in[0,1]} \times \left| s_N^{(K)}(\alpha) - \left(\frac{N(M)}{N}\right)^{(K+1)} s_{N(M)}^{(K)}(\alpha)\right|.$$

We are going to establish that this term converges to 0. For this, we first evaluate its second-order moment

$$V_2 := \mathbb{E}\left[\sup_{N(M)<N\le N(M+1)} \sup_{\alpha\in[0,1]} \times \left| s_N^{(K)}(\alpha) - \left(\frac{N(M)}{N}\right)^{(K+1)} s_{N(M)}^{(K)}(\alpha)\right|^2\right].$$

It is easy to show that

$$V_2 \le \mathbb{E}\left[\left(\frac{1}{N(M)^{K+1}} \sum_{n=N(M)}^{N(M+1)} n^K |e(n)|\right)^2\right].$$

Applying Schwartz's inequality on the RHS term of the previous inequality, it follows that

$$\begin{aligned} V_2 &\le \frac{1}{N(M)^{2(K+1)}} \left(\sum_{n=N(M)}^{N(M+1)} r_e(n,0). \sum_{n=N(M)}^{N(M+1)} n^{2K}\right) \\ &\le \mathcal{M}_2 \frac{(N(M+1)-N(M))^2}{N(M)^2} . \frac{N(M+1)^{2K}}{N(M)^{2K}}. \end{aligned}$$

Due to the definition of $N(M)$ it is easy to check that

$$\frac{(N(M+1)-N(M))^2}{N(M)^2} = \mathcal{O}\left(\frac{1}{M^2}\right) \quad \text{and} \quad \frac{N(M+1)}{N(M)} \to 1, \qquad \text{as } M\to\infty$$

which implies that

$$\mathbb{E}\left[\sup_{N(M)<N\le N(M+1)} \sup_{\alpha\in[0,1]} \left|s_N^{(K)}(\alpha) - \left(\frac{N(M)}{N}\right)^{(K+1)} s_{N(M)}^{(K)}(\alpha)\right|^2\right] = \mathcal{O}\left(\frac{1}{M^2}\right).$$

Using again Chebychev's inequality and the Borel–Cantelli lemma leads to

$$\sup_{N(M)<N\le N(M+1)} \sup_{\alpha\in[0,1]} \times \left| s_N^{(K)}(\alpha) - \left(\frac{N(M)}{N}\right)^{(K+1)} s_{N(M)}^{(K)}(\alpha)\right| \xrightarrow{\text{a.s.}} 0, \quad \text{as } M\to 0.$$

As $N(M)/N$ converges to 1 as $M\to\infty$, it follows that

$$\sup_{N(M)<N\le N(M+1)} \sup_{\alpha\in[0,1]} \left|s_N^{(K)}(\alpha) - s_{N(M)}^{(K)}(\alpha)\right| \xrightarrow{\text{a.s.}} 0, \quad \text{as } M\to 0.$$

As $\sup_{\alpha\in[0,1]} s_{N(M)}^{(K)}(\alpha)$ almost surely converges to zero, the previous equality implies that

$$\sup_{\alpha\in[0,1]} \left|s_N^{(K)}(\alpha)\right|$$

also converges to zero almost surely. □

APPENDIX II
PROOF OF ASYMPTOTIC NORMALITY OF $\boldsymbol{E}(N)$

Let $\text{cum}_L(\boldsymbol{E}(N))$ denote the Lth-order cumulant tensor of $\boldsymbol{E}(N)$. The generic form of any of its components is given by

$$\mathcal{C}_N^{(L)} := N^{-\frac{L}{2}} \sum_{n_0,\,n_1,\ldots,n_{L-1}=0}^{N-1} D^{(\nu_0)}(n_0)\cdots D^{(\nu_{L-1})}(n_{L-1}) \times \text{cum}_L\left\{e_{\tau_0}^{(\nu_0)}(n_0),\,\ldots,\,e_{\tau_{L-1}}^{(\nu_{L-1})}(n_{L-1})\right\} \tag{36}$$

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

$$D^{(0)}(n) := \begin{cases} e^{-2i\pi\alpha_0 n} \\ \frac{n}{N}\, e^{-2i\pi\alpha_0 n}, \end{cases} \quad D^{(1)}(n) := \begin{cases} e^{2i\pi\alpha_0 n} \\ \frac{n}{N}\, e^{2i\pi\alpha_0 n} \end{cases}$$

and:

$$e_\tau^{(0)}(n) := e_\tau(n), \qquad e_\tau^{(1)}(n) := \overline{e}_\tau(n).$$

It follows that $\forall\,\nu\in\{0,1\}$, and $\forall\,n\le N$, $|D^{(\nu)}(n)|\le 1$. Due to the triangular inequality applied on (36), we obtain that

$$\left|\mathcal{C}_N^{(L)}\right| \le N^{-\frac{L}{2}} \sum_{n_0,\,n_1,\ldots,n_{L-1}=0}^{N-1} \left|\text{cum}_L\left\{e_{\tau_0}^{(\nu_0)}(n_0),\, e_{\tau_1}^{(\nu_1)}(n_1),\,\ldots,\, e_{\tau_{L-1}}^{(\nu_{L-1})}(n_{L-1})\right\}\right|. \tag{37}$$

Assumption 1 implies that there is a constant $\mathcal{M}$, independent of $\{\tau_l\}_{0\le l\le L-1}$, such that $|\mathcal{C}_N^{(L)}| \le \mathcal{M}N^{-(\frac{L}{2}-1)}$. Therefore, $\text{cum}_L(\boldsymbol{E}(N)) = \mathcal{O}(N^{-(\frac{L}{2}-1)})$. If $L>2$, then $(\frac{L}{2}-1)>0$, and it follows that

$$\lim_{N\to\infty} \text{cum}_L(\boldsymbol{E}(N)) = 0$$

which implies that $\boldsymbol{E}(N)$ converges to a Gaussian distribution. □

APPENDIX III
DERIVATION OF (31) AND (32)

We first review some useful properties of fourth-order cyclic cumulants (see, e.g., [22]). Let $c_{4,\,y_a}(t,\,\boldsymbol{\tau})$ denote the fourth-order cumulant of $y_a(t)$ at time index t and lags $\boldsymbol{\tau} := [\tau_1,\,\tau_2,\,\tau_3]$

$$c_{4,\,y_a}(t,\,\boldsymbol{\tau}) := \mathrm{cum}_4\,\{y_a(t),\,y_a(t+\tau_1),\,y_a^*(t-\tau_2),\,y_a^*(t-\tau_3)\}.$$

The fourth-order cyclic cumulant at cyclic frequency α is defined by

$$c_{4,\,y_a}^{(\alpha)}(\boldsymbol{\tau}) := \int_{\mathbb{R}} c_{4,\,y_a}(t,\,\boldsymbol{\tau})e^{-2i\pi\alpha t}\,dt$$

and the corresponding cyclic trispectrum at cyclic frequency α is given by

$$S_{4,\,y_a}^{(\alpha)}(\boldsymbol{\nu}) := \int_{\mathbb{R}^3} c_{4,\,y_a}^{(\alpha)}(\boldsymbol{\tau})e^{-2i\pi\boldsymbol{\nu}\boldsymbol{\tau}^T}\,d\boldsymbol{\tau}$$

with the vector of cyclic frequencies $\boldsymbol{\nu} := [\nu_1,\,\nu_2,\,\nu_3]$. Since $y_a(t)$ is given by (1), it is well known that the cumulant cyclic frequencies are the integer multiples of $1/T_s$ and that the corresponding cyclic trispectra are given by (see, e.g., [16], [22])

$$S_{4,\,y_a}^{(k/T_s)}(\boldsymbol{\nu}) := \frac{\kappa}{T_s}\,H_a(\nu_1)H_a^*(\nu_2)H_a^*(\nu_3) \times H_a(k/T_s-\nu_1+\nu_2+\nu_3) \quad (38)$$

where $H_a(f)$ represents the Fourier transform of $h_a(t)$. As the bandwidth of $h_a(t)$ is reduced to the interval $[-(1+\rho)/2T_s,\,(1+\rho)/2T_s]$ with $0<\rho<1$, it is easy to check that the cumulant cyclic frequencies set is given by $\{k/T_s,\,|k|\le 3\}$. Therefore, the (normalized) cumulant cyclic frequencies of the digital sampled signal $y(n) := y_a(nT_e)$ are the values $\{k\alpha_0,\,|k|\le 3\}$. We denote by

$$S_{4,\,y}^{(k\alpha_0)}(\exp(2i\pi\nu_1),\,\exp(2i\pi\nu_2),\,\exp(2i\pi\nu_3))$$

the cyclic trispectrum of $y(n)$ at cyclic frequency $k\alpha_0$ and by $c_{4,\,y}^{(k\alpha_0)}(\tau_1,\,\tau_2,\,\tau_3)$ the associated cyclic cumulant sequence defined by

$$c_{4,\,y}^{(k\alpha_0)}(\tau_1,\,\tau_2,\,\tau_3) := \int_{-1/2}^{1/2} S_{4,\,y}^{(k\alpha_0)}(e^{2i\pi\nu_1},\,e^{2i\pi\nu_2},\,e^{2i\pi\nu_3})e^{2i\pi\boldsymbol{\nu}\boldsymbol{\tau}^T}\,d\boldsymbol{\nu}.$$

In the following, we will express

$$S_{4,\,y}^{(k\alpha_0)}(\exp(2i\pi\nu_1),\,\exp(2i\pi\nu_2),\,\exp(2i\pi\nu_3))$$

in terms of $S_{4,\,y_a}^{(k/T_s)}(\boldsymbol{\nu}/T_e)$. It follows that (see, e.g., [15]) for all $(\nu_1,\,\nu_2,\,\nu_3)\in(-1/2,\,1/2]^3$

$$S_{4,\,y}^{(k\alpha_0)}(e^{2i\pi\nu_1},\,e^{2i\pi\nu_2},\,e^{2i\pi\nu_3}) = \frac{1}{T_e^3}\sum_{l\in\mathbb{Z}}\sum_{\boldsymbol{\mu}\in\mathbb{Z}^3} S_{4,\,y_a}^{(l/T_s)}\left(\frac{\boldsymbol{\nu}-\boldsymbol{\mu}}{T_e}\right)\delta((k-l)\,\alpha_0 \mod 1)$$

which reduces to

$$S_{4,\,y}^{(k\alpha_0)}(e^{2i\pi\nu_1},\,e^{2i\pi\nu_2},\,e^{2i\pi\nu_3}) = \frac{1}{T_e^3}\sum_{l=-3}^{3}\sum_{\boldsymbol{\mu}\in\mathbb{Z}^3} S_{4,\,y_a}^{(l/T_s)}\left(\frac{\boldsymbol{\nu}-\boldsymbol{\mu}}{T_e}\right)\delta((k-l)\,\alpha_0 \mod 1).$$

As $h_a(t)$ is band-limited, (38) and the condition $T_e<T_s/4$ imply that

$$S_{4,\,y_a}^{(l/T_s)}(\boldsymbol{\nu}-\boldsymbol{\mu}/T_e)=0, \qquad \text{for } \boldsymbol{\mu}\neq 0$$

if $\boldsymbol{\nu}\in(-1/2,\,1/2]^3$. Hence, there is no aliasing in the RHS term of the previous equation. This leads to

$$\begin{aligned}S_{4,\,y}^{(k\alpha_0)}&(e^{2i\pi\nu_1},\,e^{2i\pi\nu_2},\,e^{2i\pi\nu_3})\\ &= \frac{1}{T_e^3}\sum_{l=-3}^{3} S_{4,\,y_a}^{(l/T_s)}\left(\frac{\boldsymbol{\nu}}{T_e}\right)\delta((k-l)\,\alpha_0 \mod 1)\\ &= \frac{1}{T_e^3}\sum_{\substack{l=-3\\ l\alpha_0=k\alpha_0 \bmod 1}}^{3} S_{4,\,y_a}^{(l/T_s)}\left(\frac{\boldsymbol{\nu}}{T_e}\right)\end{aligned}$$

for all $(\nu_1,\,\nu_2,\,\nu_3)\in(-1/2,\,1/2]^3$.

After this review on cyclic trispectra, we are able to derive (31) and (32). As $\boldsymbol{\Gamma} = S_{\boldsymbol{e}_2}^{(0)}(\exp(2i\pi\alpha_0))$, it suffices to compute the different components of previous cyclic spectra of $\boldsymbol{e}_2(n)$. Let $[\boldsymbol{U}]_{k,\,l}$ denote the $(k,\,l)$th entry of an arbitrary matrix $\boldsymbol{U}$. Using the circularity of $s(n)$, we obtain that

$$[\boldsymbol{R}_{\boldsymbol{e}_2}(n,\,\tau)]_{u,\,v} = r_y(n+v,\,\tau+u-v)r_y^*(n,\,\tau) + \mathrm{cum}_4\{y(n+u+\tau),\,y^*(n+\tau),\,y^*(n+v),\,y(n)\}$$

where $(u,\,v)\in\{-\Upsilon,\,\ldots,\,\Upsilon\}^2$. From this, we obtain that the cyclic correlation coefficients of $\boldsymbol{e}_2(n)$ at cyclic frequency 0 are given by

$$\left[\boldsymbol{R}_{\boldsymbol{e}_2}^{(0)}(\tau)\right]_{u,\,v} = \sum_{k=-1}^{1}\boldsymbol{r}_y^{(k\alpha_0)}(\tau+u-v)\boldsymbol{r}_y^{(k\alpha_0)^*}(\tau) \times e^{2i\pi k\alpha_0 v} + c_{4,\,y}^{(0)}(u+\tau,\,-\tau,\,-v).$$

Thus,

$$[\boldsymbol{\Gamma}]_{u,\,v} = \left[S_{\boldsymbol{e}_2}^{(0)}(e^{2i\pi\alpha_0})\right]_{u,\,v} = \sum_{k=-1}^{1}\mathcal{R}_{k,\,u,\,v}+\mathcal{C}_{u,\,v}$$

with

$$\mathcal{R}_{k,\,u,\,v} = \sum_{\tau\in\mathbb{Z}}\boldsymbol{r}_y^{(k\alpha_0)}(\tau+u-v)\boldsymbol{r}_y^{(k\alpha_0)^*}(\tau) \times e^{2i\pi k\alpha_0 v}e^{-2i\pi\alpha_0\tau}$$

and

$$\mathcal{C}_{u,\,v} = \sum_{\tau\in\mathbb{Z}} c_{4,\,y}^{(0)}(u+\tau,\,-\tau,\,-v)e^{-2i\pi\alpha_0\tau}.$$

Due to Parseval's identity, it follows that

$$\mathcal{R}_{k,\,u,\,v} = \int_0^1 S_y^{(k\alpha_0)}(e^{2i\pi f})\overline{S}_y^{(k\alpha_0)}\left(e^{2i\pi(f-\alpha_0)}\right) \times e^{2i\pi(u-v)f}\,df e^{2i\pi k\alpha_0 v}.$$

Since $h_a(t)$ is band-limited and $T_e<T_s/4$, the supports of the functions

$f\to S_y^{(k\alpha_0)}(\exp(2i\pi f))$ and $f\to S_y^{(k\alpha_0)}(\exp(2i\pi(f-\alpha_0)))$ are disjoint for $k=1$ and $k=-1$. Thus, $\mathcal{R}_{-1,\,u,\,v}=\mathcal{R}_{1,\,u,\,v}=0$. As for $\mathcal{C}_{u,\,v}$, it is easy to obtain that

$$\begin{aligned}\mathcal{C}_{u,\,v} &= \int_{-1/2}^{1/2} S_{4,\,y}^{(0)}(e^{2i\pi\nu_1},\,e^{2i\pi\nu_2},\,e^{2i\pi\nu_3})\\ &\quad\times\left(\sum_{\tau\in\mathbb{Z}} e^{2i\pi(\nu_1(u+\tau)-\nu_2\tau-\nu_3 v)}e^{-2i\pi\alpha_0\tau}\right)d\boldsymbol{\nu}\\ &= \int_{-1/2}^{1/2} S_{4,\,y}^{(0)}\left(e^{2i\pi\nu_1},\,e^{2i\pi(\nu_1-\alpha_0)},\,e^{2i\pi\nu_3}\right)\\ &\quad\times e^{2i\pi(\nu_1 u-\nu_3 v)}\,d\nu_1\,d\nu_3\end{aligned}$$

$$= \frac{1}{T_e^3} \int_{-1/2}^{1/2} \sum_{\substack{l=-3 \\ l\alpha_0 = 0 \bmod 1}}^{3} S_{4,\, y_a}^{(l/T_s)} \left(\frac{\nu_1}{T_e}, \frac{\nu_1 - \alpha_0}{T_e}, \frac{\nu_3}{T_e} \right) \times e^{2i\pi(\nu_1 u - \nu_3 v)}\, d\nu_1\, d\nu_3.$$

As $\alpha_0 < 1/4$, the condition ($l\alpha_0 = 0 \bmod 1$, for $|l| \le 3$) holds if and only if $l = 0$. Therefore, we have

$$\begin{aligned} \mathcal{C}_{u,\, v} &= \frac{1}{T_e^3} \int_{-1/2}^{1/2} S_{4,\, y_a}^{(0)} \left(\frac{\nu_1}{T_e}, \frac{\nu_1 - \alpha_0}{T_e}, \frac{\nu_3}{T_e} \right) \times e^{2i\pi(\nu_1 u - \nu_3 v)}\, d\nu_1\, d\nu_3 \\ &= \frac{\kappa}{T_s T_e^3} \int_{-1/2}^{1/2} H_a\left(\frac{\nu_1}{T_e}\right) H_a^*\left(\frac{\nu_1 - \alpha_0}{T_e}\right) H_a^*\left(\frac{\nu_3}{T_e}\right) \times H_a\left(\frac{\nu_3 - \alpha_0}{T_e}\right) e^{2i\pi(\nu_1 u - \nu_3 v)}\, d\nu_1\, d\nu_3. \end{aligned}$$

As $T_e < T_s/4$ (which implies that $\alpha_0 < 1/4$), $H_a(\nu_1/T_e)$ and $H_a((\nu_1 - \alpha_0)/T_e)$, respectively, are equal to $T_e h(\exp(2i\pi\nu_1))$ and $T_e h(\exp(2i\pi(\nu_1 - \alpha_0)))$ for $\nu_1 \in [-1/2,\, 1/2]$. Hence,

$$\mathcal{C}_{u,\, v} = \kappa \frac{T_e}{T_s} \int_{-1/2}^{1/2} h(e^{2i\pi\nu_1}) h^*\left(e^{2i\pi(\nu_1 - \alpha_0)}\right) h^*(e^{2i\pi\nu_3}) \times h\left(e^{2i\pi(\nu_3 - \alpha_0)}\right) e^{2i\pi(\nu_1 u - \nu_3 v)}\, d\nu_1\, d\nu_3.$$

According to (33), we obtain that

$$\begin{aligned} \mathcal{C}_{u,\, v} &= \kappa \frac{T_s}{T_e} \int_{-1/2}^{1/2} S_y^{(\alpha_0)}(e^{2i\pi f}) e^{2i\pi u\nu_1}\, d\nu_1 \int_{-1/2}^{1/2} S_y^{(\alpha_0)^*} \times (e^{2i\pi f}) e^{-2i\pi v\nu_3}\, d\nu_3 \\ &= \kappa \frac{T_s}{T_e} \boldsymbol{r}_y^{(\alpha_0)}(u) \boldsymbol{r}_y^{(\alpha_0)^*}(v). \end{aligned}$$

This proves (31). Equation (32) can be derived using similar arguments. $\square$

References

[1] K. Abed-Meraim, J. F. Cardoso, A. Y. Gorokhov, Ph. Loubaton, and E. Moulines, "On subspace methods for blind identification of single-input multiple-output FIR systems," *IEEE Trans. Signal Processing*, vol. 45, pp. 1–14, Jan. 1997.

[2] O. Besson and P. Stoica, "Frequency estimation and detection for sinusoidal signals with arbitrary envelope: A nonlinear least-squares approach," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing*, vol. IV, Seattle, WA, 1998, pp. 2209–2212.

[3] O. Besson, M. Ghogho, and A. Swami, "Parameter estimation for random amplitude chirp signals," *IEEE Trans. Signal Processing*, vol. 47, pp. 3208–3219, Dec. 1999.

[4] 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.

[5] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*. New York: Springer-Verlag, 1991.

[6] A. V. Dandawaté and G. B. Giannakis, "Statistical tests for presence of cyclostationarity," *IEEE Trans. Signal Processing*, vol. 42, pp. 2355–2369, Sept. 1994.

[7] ——, "Asymptotic theory of mixed time averages and kth-order cyclic moment and cumulant statistics," *IEEE Trans. Inform. Theory*, vol. 41, pp. 216–232, Jan. 1995.

[8] D. Dehay and H. L. Hurd, "Representation and estimation for periodically and almost periodically correlated random processes," in *Cyclostationarity in Communications and Signal Processing*. New York: IEEE Press, 1993, ch. 6.

[9] D. Dehay and J. Leskow, "Frequency determination for almost periodically correlated stochastic processes," Université de Rennes 1, France, Tech. Rep., 1994.

[10] 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.

[11] G. B. Giannakis, "Cyclostationary Signal Analysis," in *Statistical Signal Processing Section of Digital Signal Processing Handbook*, V. K. Madisetti and D. Williams, Eds. Boca Raton, FL: CRC, 1998.

[12] E. J. Hannan, "Nonlinear time series regression," *J. Appl. Probab.*, vol. 8, pp. 767–780, 1972.

[13] ——, "The estimation of frequency," *J. Appl. Probab.*, vol. 10, pp. 510–519, 1973.

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

[15] L. Izzo and A. Napolitano, "Higher-order cyclostationary properties of sampled time series," *Signal Processing*, vol. 54, no. 3, pp. 303–307, Nov. 1996.

[16] P. Marchand, "Détection et reconnaissance de modulations numériques à l'aide des statistiques cycliques d'ordre supérieur," Ph.D. dissertation, INPG-Nat. Poly. Inst. Grenoble, Grenoble, France, 1998.

[17] P. Marchand, J. L. Lacoume, and C. Le Martret, "Multiple hypothesis modulation classification based on cyclic cumulants of different orders," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 4, 1998, pp. 2157–2160.

[18] L. Mazet and Ph. Loubaton, "Cyclocorrelation based symbol rate estimation," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, vol. II, Pacific Grove, CA, Nov. 1999, pp. 1008–1012.

[19] B. Porat, *Digital Signal Processing of Random Signals*. Englewood Cliffs, NJ: Prentice Hall, 1994, Ch. 3: Parametric Estimation Theory and Ch. 4: Nonparametric Spectrum Estimation.

[20] E. Serpedin, G. B. Giannakis, A. Chevreuil, and P. Loubaton, "Blind joint estimation of carrier frequency offset and channel using nonredundant periodic modulation precoders," in *Proc. IEEE Statistical Signal and Array Processing Workshop (SSAP'98)*, Portland, OR, 1998, pp. 304–307.

[21] E. Serpedin, A. Chevreuil, G. B. Giannakis, and P. Loubaton, "Blind channel and carrier frequency offset estimation using periodic modulation procedures," *IEEE Trans. Signal Processing*, vol. 48, pp. 2389–2405, Aug. 2000.

[22] C. M. Spooner, "Higher-order statistics for nonlinear processing of cyclostationary signals," in *Cyclostationarity in Communications and Signal Processing*. New York: IEEE Press, 1993, ch. 2.

[23] 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.