



HAL
open science

Time average estimation in the fraction-of-time probability framework

Dominique Dehay, Jacek Leśkow, Antonio Napolitano

► **To cite this version:**

Dominique Dehay, Jacek Leśkow, Antonio Napolitano. Time average estimation in the fraction-of-time probability framework. *Signal Processing*, 2018, 153, pp.275-290. <10.1016/j.sigpro.2018.07.005>. <hal-01862710>

HAL Id: hal-01862710

<https://hal.science/hal-01862710v1>

Submitted on 28 Aug 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Time Average Estimation in the Fraction-of-Time Probability Framework

Dominique Dehay

Univ. Rennes, CNRS, IRMAR – UMR 6625, Université Rennes 2, F-35000, Rennes, France.

Jacek Leśkow

Cracow University of Technology, Cracow, Poland.

Antonio Napolitano*

EURASIP Member, University of Napoli “Parthenope”, Department of Engineering, 80143, Napoli, Italy

Abstract

The problem of time average estimation is addressed in the fraction-of-time probability framework. In this approach, the observed signal is modeled as a single function of time rather than as a sample path of a stochastic process. Under mild regularity assumptions on temporal cumulants of the signal, a central limit theorem (CLT) is proved for the normalized error of the time average estimate, where the normalizing factor is the square root of the observation interval length. This rate of convergence is the same as that obtained in the classical stochastic approach, but is derived here without resorting to mixing assumptions. That is, no ergodicity hypothesis is needed. Examples of functions of interest in communications are provided whose time average estimate satisfies the CLT. For the class of the almost-periodic functions, a non normal limit distribution for the normalized error of the time average estimate can be obtained when the normalizing factor equals the observation interval length. The limit distribution, nevertheless, depends on the sequence adopted to perform the limit, a result that cannot be obtained in the classical stochastic approach. Numerical examples illustrate the theoretical results and an application to test the presence of a nonzero-mean cyclostationary signal is presented.

Keywords: Time average; Central limit theorem; Fraction-of-time probability; Functional approach; Pseudo-random functions; Almost-periodic functions.

1. Introduction

In the classical stochastic approach, signals or time series are modeled as realizations of stochastic processes [14]. For this purpose, a sample space Ω equipped with a σ -field \mathcal{F} and a probability measure P defined on \mathcal{F} are introduced. Probabilistic functions are defined as integrals of functions of the stochastic processes made with respect to the probability measure. Estimates of these probabilistic functions are time averages, over finite time intervals, of functions of a single realization. Then some mixing assumptions assure the convergence of the estimates

*Corresponding author

Email addresses: dominique.dehay@univ-rennes2.fr (Dominique Dehay), jleskow@pk.edu.pl (Jacek Leśkow), antonio.napolitano@uniparthenope.it antnapol@gmail.com (Antonio Napolitano)

to the corresponding probabilistic functions as the time interval becomes infinite (ergodicity theorems [25]).

In practice, it is difficult if not impossible to verify the mixing assumptions. In addition, quite frequently in practical cases only one realization of the time series exists. This is the case, for example, of a biological signal measured for a specific living being, the signal coming from a specific astronomic source, a communication signal arising from a specific transmitter, and the exchange rate between two specific currencies.

An alternative approach for signal analysis consists in modeling the observed signal or time series as a single function of time. Such a signal is modeled as a segment of an idealized one existing on the whole real line (that is, a “persistent” signal). Such an approach is referred to as *functional approach* [5, Sec. 2.11] or *fraction-of-time (FOT) probability approach* [17, Chap. 8] and was first adopted in [44] for correlation and spectral analysis (see [29] for an extensive historical perspective).

The key concept in the FOT approach is the relative measure μ_R of sets introduced in [27]. For a Borel set A , its relative measure $\mu_R(A)$ is defined as the Lebesgue measure of $A \cap [-T/2, T/2]$, divided by T , and considered in the limit as $T \rightarrow \infty$. Such a measure plays an equivalent role to the theoretical probability P in the stochastic approach where the object (Ω, \mathcal{F}, P) is considered. Once the concept of relative measurability is introduced for a signal $x(t)$, a valid distribution function can be constructed whose expectation is the infinite time average [29]. The distribution at point ξ represents the fraction-of-time that the signal is below the level ξ [17, Chap. 8]. Using such distribution and expectation, all the familiar probabilistic functions as mean, moments and cumulants can be constructed. Moreover, considering two or more time-shifted versions of the same signal, concepts can be extended to the multivariate case leading to the definition of joint distributions, autocorrelation, cross moments, and cumulants [18], [19], [29]. Independence of functions in the functional approach can be defined [27], [26], and a central limit theorem (CLT) proved [12]. Higher-order statistics are considered in [20], [39] and a link with the relative measure is established in [34, Chap. 6]. Further results on signals modeled as single functions of time are in [10], [24], [31], [32].

The classical probability measure is σ -additive and continuous and, hence, the ensemble expectation is σ -linear. These properties are a consequence of the implicit assumption in the theory of probability that the sample space Ω has a finite measure [21], [22]. Thus, the normalization $P(\Omega) = 1$ is obtained just dividing a measure by the (finite) measure of the whole space. In contrast, the analogous normalization of the relative measure $\mu_R(\mathbb{R}) = 1$ is obtained by a limit process as $T \rightarrow \infty$. Consequently, the relative measure is not σ -additive and not continuous. Moreover, the infinite time average, unlike the ensemble expectation, is not σ -linear [29]. Such lack of σ -additivity and σ -linearity is not an obstacle using the FOT approach in practice when dealing with time averages. It is, rather, a motivation. In fact, assumptions on the stochastic process and consequent results expressed in terms of ensemble averages could not be explained in terms of time averages as needed by the practitioner. In contrast, in the FOT approach, assumptions are made directly on the unique available time series at hand and results are proved by proving the convergence of time averages when the observation interval grows to infinity.

In this paper, the problem of infinite time average estimation of a continuous-time persistent signal is addressed in the fraction-of-time probability framework. Such problem is fundamental in signal processing. For example, time average estimation occurs in the following problems: estimation of any coefficient of a Fourier series [4, chap. 1], cyclic autocorrelation estimation by

the cyclic correlogram [18, Chap. 11, Sec. A], cyclic spectrum estimation by the time-smoothed cyclic periodogram [18, Chap. 13, Sec. A], characteristic function, moment, and cumulant estimation.

In both stochastic and FOT approaches, the estimator is the time average of the signal over the finite time interval $[t - T/2, t + T/2]$ centered in t and with width T . In the stochastic approach, t is fixed (and generally assumed to be 0 or $T/2$) and the variability of the estimate is due to the sample path chosen for the estimation, that is, the estimate is a random variable (it depends on $\omega \in \Omega$). Under appropriate mixing and stationarity assumptions, as $T \rightarrow \infty$ the estimate converges in some probabilistic sense to the expected value of the process. In contrast, in the FOT approach, the variability of the estimate is due to t , the central point of the observation interval, when t ranges in a wider temporal interval, say $[-Z/2, Z/2]$, with $Z \gg T$ [34, Sec. 6.3.5]. In the FOT approach, asymptotic results are expressed in terms of a double limit as $Z \rightarrow \infty$ (to average the variability with respect to t , the central point of the observation interval) and as $T \rightarrow \infty$ (to consider the whole time series), provided that $Z/T \rightarrow \infty$.

In a practical situation where a unique signal source exists, the experimenter observes different segments of the unique signal emitted or generated by the source, computes an estimate of the time average for each segment, and then averages these estimates to reduce variability. Moreover, the estimates computed on different segments can be used to compute the empirical distribution of the error and possibly the confidence intervals. This is exactly the model of the estimation theory in the FOT approach.

Under assumptions of summability of FOT cumulants, a CLT is proved for \sqrt{T} times the estimation error, which is the same rate of convergence found in the classical stochastic approach. That is, it is shown that the normalized error, as a function of time, has a normal distribution when the data-record length T becomes infinitely large. It is worth emphasizing that the assumptions are given here in terms of properties (summability of FOT cumulants) of the single function of time. Such assumptions can be analytically verified. In contrast, analogous CLT theorems in the classical stochastic approach make mixing assumptions [11], possibly expressed in terms of summability of cumulants of stochastic processes, on the ensemble of realizations that constitute the stochastic process. In particular, in the classical stochastic approach the analysis of rates of convergence is generally made under ergodicity conditions. The proposed approach does not require ergodicity hypothesis and is alternative to different approaches that have been proposed when the observed time series is not ergodic [9]. As numerical illustrative example on the CLT, a Bass function is analyzed [3, Secs. II and V-3-5]. It describes a sample path of a binary pulse-amplitude modulated (PAM) signal with stationary white modulating sequence.

Note that the cumulant summability assumption is not verified in the case of long range dependence [37, Chap. III] occurring for impulsive noise encountered in several application fields [16], [40], [41], [42]. In such case, the proposed approach cannot be applied and further analysis is required.

The almost-periodic functions are signals of interest in many applications [2], [4], [18], [34], [35] that do not satisfy the assumptions for the CLT. Also in this case cumulants are not summable. However, in the paper it is shown that T times the estimation error has non zero bounded covariance and cumulants and convergence of estimates of covariance, cumulants, and distribution of the normalized estimation error can be obtained only along appropriate sequences $\{T_n\}$ of data-record lengths. Such a result does not have a counterpart in the classical stochastic

approach and emphasizes the usefulness of the FOT approach as an alternative.

As an example of application, the problem of estimating the cyclic autocorrelation function of a non-zero mean cyclostationary signal is addressed in the FOT probability framework. A result that cannot be established in the classical stochastic approach is found. Then, the FOT counterpart of a statistical test for presence of cyclostationarity originally presented in the stochastic approach [8] is shown to be effective even in the case of strong almost periodic component superimposed to the cyclostationary signal of interest.

The paper is organized as follows. In Section 2, a brief overview of the FOT probability approach is provided to introduce notation and definitions. In Section 3, the problem of time average estimation is addressed and a CLT theorem is stated for the normalized error. Pseudo-random functions and almost-periodic functions are treated in detail in Sections 4 and 5, respectively. Numerical results are reported in Section 6 and an example of application in Section 7. Conclusions are drawn in Section 8. Proofs are reported in Appendix.

2. Fraction-of-Time Probability Framework for Signal Analysis

In this section, the fraction-of-time (or functional) approach for signal analysis is briefly reviewed in order to introduce definitions and notation. See [12], [29] for further details.

Definition 2.1. *Let us consider the set $A \in \mathcal{B}_{\mathbb{R}}$, where $\mathcal{B}_{\mathbb{R}}$ is the σ -field of the Borel subsets and μ is the Lebesgue measure on the real line \mathbb{R} . The relative measure of A is defined as [27]*

$$\mu_R(A) \triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \mu(A \cap [t_0 - T/2, t_0 + T/2]) \quad (2.1)$$

provided that the limit exists. In such a case, the limit does not depend on t_0 and the set A is said to be relatively measurable (RM). \square

The relative measure is additive [29, Fact 2.4] but not σ -additive [29, Fact 2.5].

Definition 2.2. *Let $x(t)$ be a Lebesgue measurable function. The function $x(t)$ is said to be relatively measurable if and only if the set $\{t \in \mathbb{R} : x(t) \leq \xi\}$ is RM for every $\xi \in \mathbb{R} - \Xi_0$, where Ξ_0 is at most a countable set of points. \square*

Examples of RM functions are the almost-periodic functions and some pseudo-random functions [29, Sec. 6]. Functions that are not RM can be easily constructed, in contrast to functions that are not Lebesgue measurable [29], [30].

Let $\mathbf{1}_A$ be the indicator function of the set A , that is, $\mathbf{1}_A = 1$ if $t \in A$ and $\mathbf{1}_A = 0$ if $t \notin A$ and let us denote by

$$\langle \phi(t) \rangle_t \equiv \lim_{T \rightarrow \infty} \frac{1}{T} \int_{t_0 - T/2}^{t_0 + T/2} \phi(t) dt \quad (2.2)$$

the infinite time average of the function $\phi(t)$, where the limit is independent of t_0 .

Definition 2.3. *Each RM function $x(t)$ generates a function*

$$F_x(\xi) \triangleq \mu_R(\{t \in \mathbb{R} : x(t) \leq \xi\}) = \langle \mathbf{1}_{\{t \in \mathbb{R} : x(t) \leq \xi\}} \rangle_t \quad (2.3)$$

in all points ξ where the limit (2.1) exists.

The function $F_x(\xi)$ is called FOT distribution of $x(t)$. It has all the properties of a valid distribution function except for the right continuity in the discontinuity points [29]. \square

The function $F_x(\xi)$ is obtained in the limit as $T \rightarrow \infty$ of the finite-time FOT distribution

$$F_{T,t}(\xi) \triangleq \frac{1}{T} \mu(\{u \in [t - T/2, t + T/2] : x(u) \leq \xi\}) \quad (2.4)$$

which has all the properties of a valid distribution function including the right continuity in the discontinuity points [29]. The expectation induced by $F_{T,t}(\xi)$ is the finite time average on the interval $[t - T/2, t + T/2]$.

The expectation induced by the FOT distribution (2.3) is the infinite time average (2.2) [29, Corollary 3.1] (stationary FOT model [34, Sec. 6.2]). It is not σ -linear.

The relative measure μ_R plays an equivalent role to the theoretical probability P in the stochastic approach where the object (Ω, \mathcal{F}, P) is considered [29]. Similarly, the infinite time average plays, in the FOT probability framework, the same role played by the ensemble operator $E\{\cdot\}$ in the classical stochastic framework. For example, for a stationary stochastic process $X(t)$ we have that the cumulative distribution function can be expressed as

$$F_X(\xi) = E\{\mathbf{1}_{\{X(t) \leq \xi\}}\} \quad (2.5)$$

which is the stochastic counterpart of (2.3).

Definition 2.4. *Let $x(t)$ be a RM function. The FOT characteristic function $\Phi_x(\omega)$ of $x(t)$ is defined as*

$$\Phi_x(\omega) \triangleq \left\langle e^{j\omega x(t)} \right\rangle_t = \int_{\mathbb{R}} e^{j\omega \xi} dF_x(\xi) \quad (2.6)$$

where the second equality is consequence of the fundamental theorem of expectation in the functional approach [29, Theorem 3.2]. \square

Let Ξ_0 be at most a countable set of $(k - 1)$ -dimensional manifolds of \mathbb{R}^k . The joint characterization of k signals is made starting from the following definition.

Definition 2.5. *The Lebesgue measurable functions $x_1(t), \dots, x_k(t)$ are said to be jointly relatively measurable if the limit*

$$\begin{aligned} F_{x_1 \dots x_k}(\xi_1, \dots, \xi_k) &\triangleq \mu_R(\{t \in \mathbb{R} : x_1(t) \leq \xi_1\} \cap \dots \cap \{t \in \mathbb{R} : x_k(t) \leq \xi_k\}) \\ &= \left\langle \mathbf{1}_{\{t \in \mathbb{R} : x_1(t) \leq \xi_1\}} \cdots \mathbf{1}_{\{t \in \mathbb{R} : x_k(t) \leq \xi_k\}} \right\rangle_t \end{aligned} \quad (2.7)$$

exists for all $(\xi_1, \dots, \xi_k) \in \mathbb{R}^k - \Xi_0$. \square

In most cases, the joint relative measurability of any time-shifted versions of signals is required. This is the case, for example, when the cross-correlation and cross-moments of two or more signals must be defined [29, Corollary 4.2].

The independence of sets is defined in terms of relative measure in [12, Def. 2.4], [27]. Then, the independence of functions is defined in [12, Def. 2.5], [27], and the following result is proved.

Theorem 2.6. [12, Theorem 2.6]. *The jointly RM functions $x_1(t), \dots, x_k(t)$ are FOT-independent if and only if*

$$F_{x_1 \dots x_k}(\xi_1, \dots, \xi_k) = F_{x_1}(\xi_1) \cdots F_{x_k}(\xi_k) \quad \forall (\xi_1, \dots, \xi_k) \in \mathbb{R}^k - \Xi_0. \quad (2.8)$$

Note that in the FOT approach the independence condition (2.8) is a theorem. In contrast, in the classical stochastic approach, the factorization of the k th-order joint distribution into the product of the k marginal distributions is assumed as definition of independence for k stochastic processes [14, Sec. 1.3].

Definition 2.7. Let $x_1(t), \dots, x_k(t)$ be jointly RM functions. Their joint cumulant is defined as

$$\text{cum}_t \{x_1(t), \dots, x_k(t)\} \triangleq (-j)^k \frac{\partial^k}{\partial \omega_1 \dots \partial \omega_k} \log \left\langle e^{j\omega_1 x_1(t)} \dots e^{j\omega_k x_k(t)} \right\rangle_t \Big|_{\omega=0} \quad (2.9)$$

where $\omega \triangleq [\omega_1, \dots, \omega_k]$. For $k = 2$ we have the FOT covariance $\text{cov}_t[x_1(t+s), x_2(t)] \triangleq \langle x_1(t+s) x_2(t) \rangle_t - \langle x_1(t) \rangle_t \langle x_2(t) \rangle_t$. \square

Definition 2.8. Let $\{x_n(t)\}_{n \in \mathbb{N}}$ be a sequence of RM functions. If a RM function $x(t)$ exists such that, for every real a and b ,

$$\lim_{n \rightarrow \infty} \mu_R \left(\{t \in \mathbb{R} : a < x_n(t) \leq b\} \right) = \mu_R \left(\{t \in \mathbb{R} : a < x(t) \leq b\} \right) \quad (2.10)$$

the sequence $\{x_n(t)\}_{n \in \mathbb{N}}$ is said to be convergent with respect to μ_R to $x(t)$ (RM weak convergence or convergence in FOT distribution). \square

Sufficient conditions for weak convergence are given in the assumptions of [12, Theorem 3.1].

Definition 2.8 can be straightforwardly extended to define the weak convergence (in terms of joint relative measure) of jointly RM functions.

3. Time Average Estimation

In this section, for the time average estimate, the FOT characterization is derived in terms of FOT bias, covariance, and distribution. For this purpose, assumptions are made directly on the observed signal, without invoking the existence of an underlying stochastic process whose signal at hand is a realization.

Let $x(t)$ be a bounded RM function. Its time average is defined as

$$m \triangleq \langle x(t) \rangle_t = \int_{\mathbb{R}} \xi \, dF_x(\xi) \quad (3.1)$$

where the second equality is due to [29, Theorem 3.2].

Definition 3.1. Let $x(t)$ be a RM function with time average m . The estimator $\hat{m}_T(t)$ of m based on the observation $x(u)$, $u \in [t - T/2, t + T/2]$, is defined as

$$\hat{m}_T(t) \triangleq \frac{1}{T} \int_{t-T/2}^{t+T/2} x(u) \, du = \frac{1}{T} \int_{-T/2}^{T/2} x(u+t) \, du = \int_{\mathbb{R}} \xi \, dF_{T,t}(\xi). \quad (3.2)$$

Note that in the classical stochastic approach, $x(t)$ is modeled as a sample path of a stochastic process, that is $x(t) = X(t, \omega)$ and the estimator of the time average has the same expression (3.2). However, in such a case t is considered fixed (and generally assumed to be 0 or $T/2$) and $\hat{m}_T(t)$ is a random variable. That is, its variability is due to the dependence on $\omega \in \Omega$. If some mixing and stationarity assumptions hold, as $T \rightarrow \infty$ the estimate converges in some

probabilistic sense to the expected value $E\{X(t, \omega)\}$ of the process [5], [14, Chap. XI], [21], [22], [37, Chap. III]. In contrast, in the FOT approach, the variability of the estimate is due to t , the central point of the observation interval, when t ranges in a wider temporal interval, say $[-Z/2, Z/2]$, with $Z \gg T$ [34, Sec. 6.3.5].

Theorem 3.2. *Let $x(t)$ be a bounded RM function with time average m . It results that*

$$\text{bias}[\widehat{m}_T(t)] \triangleq \lim_{Z \rightarrow \infty} \frac{1}{Z} \int_{-Z/2}^{Z/2} [\widehat{m}_T(t) - m] dt = 0. \quad (3.3)$$

Proof: *See Appendix.* □

From Theorem 3.2 it follows that the rate of convergence as $T \rightarrow \infty$ of the bias of estimator $\widehat{m}_T(t)$ is arbitrarily fast for every bounded RM function (in particular for pseudo-random functions (Sec. 4) and almost-periodic functions (Sec. 5)). That is, we have

$$\lim_{T \rightarrow \infty} T^r \text{bias}[\widehat{m}_T(t)] = 0 \quad \forall r > 0. \quad (3.4)$$

3.1. Normalized Time Average

In this section, a FOT statistical characterization of the estimation error for the time average is provided. The obtained results are useful in deriving confidence intervals for the estimate.

Assumption 3.3. *For any $k \geq 2$ and any $(\tau_1, \dots, \tau_k) \in \mathbb{R}^k$ the translated time series $x(t + \tau_1), \dots, x(t + \tau_k)$ are jointly RM.* □

Since the translate of an almost-periodic function is in turn almost-periodic, and almost-periodic functions are jointly RM [29, Theorem 6.3], we have that Assumption 3.3 is verified by the translates of almost periodic functions. In addition, this assumption is verified by the translates of a Agnew-Kac function (Sec. 4.2). Note that, in general, the translates of a RM function are not necessarily jointly RM [29, Example 6.1].

Assumption 3.4. *The function $x(t)$ has a summable FOT covariance. That is*

$$\int_{\mathbb{R}} |\text{cov}_t[x(t+s), x(t)]| ds < \infty. \quad (3.5)$$

Assumption 3.5. *For any $k \geq 2$ and any $(\tau_1, \dots, \tau_{k-1}) \in \mathbb{R}^{k-1}$ the functions $x(t + \tau_1), \dots, x(t + \tau_{k-1}), x(t)$ have joint k th-order cumulant such that, when $T \rightarrow \infty$*

$$\underbrace{\int_{u-T/2}^{u+T/2} \cdots \int_{u-T/2}^{u+T/2}}_{k-1} |\text{cum}_t[x(t + \tau_1 + s_1), \dots, x(t + \tau_{k-1} + s_{k-1}), x(t)]| ds_1 \cdots ds_{k-1} = o(T^{k/2-1}) \quad (3.6)$$

where $o(\cdot)$ can depend on $\tau_1, \dots, \tau_{k-1}$. In (3.6), we can put $u = 0$ since the condition holds as $T \rightarrow \infty$. □

Note that Assumption 3.5 is verified if the functions $x(t + \tau_1), \dots, x(t + \tau_{k-1}), x(t)$ have summable joint k th-order cumulant:

$$\int_{\mathbb{R}^{k-1}} |\text{cum}_t [x(t + \tau_1 + s_1), \dots, x(t + \tau_{k-1} + s_{k-1}), x(t)]| ds_1 \cdots ds_{k-1} < \infty. \quad (3.7)$$

Assumptions 3.4 and 3.5 are verified if the signal $x(t)$ and its time-shifted version $x(t + \tau)$ are approximately FOT independent (see Theorem 2.6) when $|\tau|$ becomes large.

It is important to emphasize that Assumptions 3.3–3.5 refer to the single available function of time. They do not involve any underlying stochastic process and their validity can be directly verified on this single function.

Definition 3.6. Let $x(t)$ be a bounded RM function. The normalized time average with rate of convergence T^a , $a > 0$, is defined as

$$\tilde{m}_T(t) \triangleq T^a [\hat{m}_T(t) - m]. \quad (3.8)$$

Theorem 3.7. Central Limit Theorem. Under Assumptions 3.4 and 3.5, for any $k \geq 2$ and any $(\tau_1, \dots, \tau_k) \in \mathbb{R}^k$, the vector of normalized time averages $[\tilde{m}_T(t + \tau_1), \dots, \tilde{m}_T(t + \tau_k)]$ having each entry defined as in (3.8) with $a = 1/2$, converges in FOT-distribution to the multidimensional Gaussian law with mean vector $[0, \dots, 0]$ and covariance matrix $\{K(\tau_i - \tau_j)\}_{i,j=1}^k$ defined by

$$K(\tau) \triangleq \lim_{T \rightarrow \infty} \text{cov}_t [\tilde{m}_T(t + \tau), \tilde{m}_T(t)] = \int_{\mathbb{R}} \text{cov}_t [x(t + \tau + s), x(t)] ds. \quad (3.9)$$

Proof: See Appendix. □

Note that Theorem 3.7 is different from the CLT in [12]. In fact, in [12] it is proved that a properly normalized sum of functions independent in the FOT sense has a normal FOT distribution when the number of functions tends to infinity. In contrast, only one function is considered in Theorem 3.7.

The well-known property that the marginal distributions obtained from a multidimensional Gaussian distribution are in turn Gaussian leads to the following result which is an immediate consequence of Theorem 3.7.

Corollary 3.8. Under Assumptions 3.4 and 3.5, the normalized time average $\tilde{m}_T(t)$ defined in (3.8) with $a = 1/2$ converges in FOT-distribution to the Gaussian law with zero mean and variance $K(0)$. Hence, when $K(0) > 0$ we have

$$\lim_{T \rightarrow \infty} F_{\tilde{m}_T}(\xi) = \lim_{T \rightarrow \infty} \mu_R \left(\{t \in \mathbb{R} : \tilde{m}_T(t) \leq \xi\} \right) = \frac{1}{\sqrt{2\pi K(0)}} \int_{-\infty}^{\xi} e^{-s^2/2K(0)} ds. \quad (3.10)$$

□

Once the FOT distribution is obtained for the estimate, confidence intervals can be defined exactly as in the classical stochastic approach.

4. Pseudo-Random Functions

In this section, pseudo-random functions are considered. These functions are deterministic with erratic behavior. At first, pseudo-random functions with ideal characteristics such that the CLT (Theorem 3.7) is satisfied are defined and their FOT characterization provided. Then, practical examples of functions that approximately fit this characterization are considered. Specifically, the Bass pseudo-random functions [3, Section V-3-5] and PAM signals with modulating sequences obtained by a linear congruential generator (LCG) [23] are considered.

4.1. Ideal Pseudo-Random Functions

Definition 4.1. A deterministic sequence w_n , $n \in \mathbb{Z}$, is said to be pseudo-random white, in short white, if it is zero mean and, for every integer $k \geq 2$, the k th-order cumulant of its shifted versions is given by

$$\begin{aligned} \text{cum}_n \{w_{n+m_1}, \dots, w_{n+m_k}\} &\triangleq (-j)^k \frac{\partial^k}{\partial \omega_1 \dots \partial \omega_k} \log \langle e^{j\omega_1 w_{n+m_1}} \dots e^{j\omega_k w_{n+m_k}} \rangle_n \Big|_{\boldsymbol{\omega}=\mathbf{0}} \\ &= c_k \delta_{\mathbf{m}'-\mathbf{1}m_k} \end{aligned} \quad (4.1)$$

where $\langle \cdot \rangle_n$ denotes the discrete-time infinite time average, $\mathbf{m}' \triangleq [m_1, \dots, m_{k-1}]$, $\mathbf{1} \triangleq [1, \dots, 1]$, and δ_γ is the Kronecker delta, that is, $\delta_\gamma = 1$ for $\gamma = 0$ and $\delta_\gamma = 0$ for $\gamma \neq 0$. Therefore $\delta_{\mathbf{m}'-\mathbf{1}m_k} = 1$ only if $m_1 = \dots = m_{k-1} = m_k$. In (4.1), c_k is the k th-order FOT cumulant of the sequence w_n defined as

$$c_k \triangleq (-j)^k \frac{d^k}{d\omega^k} \langle e^{j\omega w_n} \rangle_n \Big|_{\omega=0}. \quad (4.2)$$

□

The behavior of such a sequence is unpredictable. For this reason w_n is referred to as a pseudo-random sequence.

From Definition 4.1, it follows that the pseudo-random white sequence w_n can be seen as a typical (i.e., satisfying the weak law of large numbers) realization of a sequence of independent and identically distributed (i.i.d.) random variables.

Definition 4.2. Let $\{w_n\}_{n \in \mathbb{Z}}$ be the pseudo-random white sequence of Definition 4.1. The ideal pseudo-random function $x(t)$ is defined as

$$x(t) \triangleq \sum_{n=-\infty}^{+\infty} w_n q(t - nT_p) \quad t \in \mathbb{R} \quad (4.3)$$

provided that the series is convergent. In (4.3), T_p is the symbol period and $q(t)$ is bounded and $q(t) \in L^1(\mathbb{R})$. □

If $q(t)$ has finite support, then a finite number of terms are non zero in the right-hand side of (4.3) and the series is absolutely convergent. If $q(t)$ has not necessarily finite support and the sequence w_n is bounded, then $|w_n| \leq M$, for some $M > 0$, $\forall n$, and we have

$$\left| \sum_{n=-\infty}^{+\infty} w_n q(t - nT_p) \right| \leq M \sum_{n=-\infty}^{+\infty} |q(t - nT_p)|.$$

In such a case, convergence in the right-hand side of (4.3) is assured by the condition $q(t) = \mathcal{O}(|t|^{-2})$ as $|t| \rightarrow \infty$.

Lemma 4.3. *The pseudo-random function defined in (4.3) has k th-order cumulant*

$$\text{cum}_t \{x(t + \tau_1), \dots, x(t + \tau_k)\} = \frac{c_k}{T_p} \int_{\mathbb{R}} q(t) \prod_{i=1}^{k-1} q(t + \tau_i - \tau_k) dt. \quad (4.4)$$

Proof: *See Appendix.* □

Theorem 4.4. *Let $x(t)$ be the pseudo-random function defined in (4.3) and $\tilde{m}_T(t)$ its normalized time average with rate of convergence T^a defined in (3.8). The k th-order cumulant of $\tilde{m}_T(t + \tau_1), \dots, \tilde{m}_T(t + \tau_k)$ is such that*

$$\text{cum}_t \{\tilde{m}_T(t + \tau_1), \dots, \tilde{m}_T(t + \tau_k)\} = \mathcal{O}(T^{k(a-1)+1}). \quad (4.5)$$

For $k = 2$ and $a = 1/2$ we have

$$\lim_{T \rightarrow \infty} \text{cov}_t \{\tilde{m}_T(t + \tau), \tilde{m}_T(t)\} = \frac{c_2}{T_p} \left[\int_{\mathbb{R}} q(t) dt \right]^2. \quad (4.6)$$

Proof: *See Appendix.* □

Theorem 4.5. *Under the assumptions of Theorem 4.4, for $a = 1/2$, the functions $\tilde{m}_T(t + \tau_1), \dots, \tilde{m}_T(t + \tau_k)$ are asymptotically ($T \rightarrow \infty$) jointly normal.*

Proof: *See Appendix.* □

4.2. Agnew-Kac Functions

Definition 4.6. *The Agnew-Kac sequence is defined as [1]*

$$a_n \triangleq \cos(e^{n^2}) \quad n \in \mathbb{Z}. \quad (4.7)$$

Theorem 4.7. *k different shifted versions of the Agnew-Kac sequence $a_{n+m_1} = \cos(e^{(n+m_1)^2}), \dots, a_{n+m_k} = \cos(e^{(n+m_k)^2})$, with $m_i \neq m_j$ for $i \neq j$, are independent in the FOT sense [12, Definition 2.5, Theorem 2.6], [29, Sec. 5].*

Proof: *See Appendix.* □

Corollary 4.8. *The Agnew-Kac sequence defined in (4.7) is white. That is, for every $k \geq 2$*

$$\text{cum}_n \{a_{n+m_1}, \dots, a_{n+m_k}\} = c_k \delta_{\mathbf{m}' - \mathbf{1}_{m_k}}. \quad (4.8)$$

Proof: *See Appendix.* □

For $k = 2$, (4.8) can be expressed as $\langle a_{n+m_1} a_{n+m_2} \rangle_n = \delta_{m_1 - m_2}$, which is the autocorrelation of the Agnew-Kac sequence a_n .

Definition 4.9. *The Agnew-Kac pseudo-random function is defined as the pseudo-random function (4.3) with sequence w_n replaced by the Agnew-Kac sequence defined in (4.7).* □

From Corollary 4.8 it follows that the Agnew-Kac sequence defined in (4.7) is a pseudo-random white sequence according to Definition 4.1. Thus the Agnew-Kac pseudo-random function of Definition 4.9 is an example of ideal pseudo-random function (Definition 4.2). Consequently, for the Agnew-Kac pseudo-random function Lemma 4.3, Theorem 4.4, and Theorem 4.5 hold.

The whiteness property of the Agnew-Kac sequence is a consequence of the very rapid variability of the argument of cosine in (4.7). Very rapidly increasing sequences, other than e^{n^2} , can be used in the argument of cosine to obtain other examples of pseudo-random white sequences.

From a computational point of view, note that the numerical values of e^{n^2} and similar rapidly increasing sequences, when stored in a (finite-length) memory register, are truncated even for very small values of $|n|$. For example, in a double precision register, the value of e^{n^2} is truncated even for $|n| < 10$. Thus, the Agnew-Kac sequence and similar sequences cannot be suitably adopted to generate a white sequence in computer experiments. For this reason, sequences less rapidly varying should be used as argument of cosine. This simplifies numerical problems but gives rise to sequences whose whiteness is only approximate [1].

In the Sections 4.3 and 4.4, two examples of classes of binary pseudo-random sequences that closely approximate a white pseudo-random sequence are presented.

4.3. Bass Functions

Definition 4.10. Let $P(\cdot)$ be a polynomial of degree greater than or equal to two with at least one irrational coefficient for the second or higher power. The Bass pseudo-random sequence [3, Section V-3-5] is defined as

$$b_n \triangleq \cos(\pi \lfloor P(n) \rfloor) = e^{j\pi \lfloor P(n) \rfloor} \quad n \in \mathbb{Z} \quad (4.9)$$

where $\lfloor \cdot \rfloor$ denotes the floor operator. □

The sequence b_n assumes values ± 1 with the same FOT probability. Thus, its time average is zero.

Definition 4.11. Let $\{b_n\}_{n \in \mathbb{Z}}$ be the Bass pseudo-random sequence of Definition 4.10. The Bass pseudo-random function [3, Section V-3-5] is defined as $x(t)$ in (4.3) with w_n replaced by b_n . □

Simulation results have shown that cumulants of $b_{n+m_1}, \dots, b_{n+m_k}$, as functions of m_1, \dots, m_k are not spread and closely approach the expression in (4.1) up to a given order k , provided that the order of the polynomial $P(\cdot)$ in definition 4.10 is sufficiently high and the coefficient of the highest order power is irrational. In such a case, one obtains a close approximation of the normal law for $\tilde{m}_T(t)$ with $a = 1/2$ (Section 6.1).

4.4. PAM Signal with Modulating Sequence Obtained by a LCG

The *linear congruent generator* is defined by the recurrence relation

$$x_{n+1} = (a x_n + c) \bmod m. \quad (4.10)$$

The parameters m , a , and c are integers. m is called *modulus* and $m > 0$; a is called *multiplier* and $0 < a < m$; c is called *increment* and $0 \leq c < m$; x_0 is called *seed* and $0 \leq x_0 < m$. All values assumed by integers x_n are less than m and the sequence x_n is periodic. The period of a general mixed congruential generator depends on the choice of the parameters m , a , and c . It is at most m and in some cases it can be much less than m .

Theorem 4.12 ([23]). *The sequence defined in (4.10) has full period m for all seed values provided that 1) m and c are relatively prime; 2) $a - 1$ is divisible by all prime factors of m ; 3) $a - 1$ is divisible by 4 if m is divisible by 4. \square*

Different finite length sequences $\{x_n, n = n_0, \dots, n_0 + N - 1\}$ are obtained by initializing (4.10) with different seeds. Each sequence, corresponding to a given n_0 , is a segment of the unique periodic sequence defined in (4.10) with very large period. However, according to the FOT model, each $\{x_n, n = n_0, \dots, n_0 + N - 1\}$ can be seen as a segment of an ideally infinitely long aperiodic sequence, a pseudo-random sequence, provided that the overall observation interval is smaller than the period of the periodic sequence defined in (4.10). This is the case in most applications.

A pseudo-random sequence of numbers in the interval $[0, 1)$ is obtained by taking the sequence $u_n \triangleq x_n/m$. Thus, $v_n \triangleq \text{sign}(u_n - 1/2)$ is a binary pseudo-random sequence. A pseudo-random function (binary PAM signal) is then obtained by replacing in (4.3) the sequence w_n with v_n .

Several pseudo-random generators in various software packages are based on the LCG or its variants. Statistical tests on these sequences are investigated, for example, in [7], [15].

5. Almost-Periodic Functions

In this section, almost-periodic functions [4, chap. 1] are defined and their FOT characterization provided.

5.1. FOT Characterization

Definition 5.1. *A function $x(t)$ is said to be uniformly almost-periodic (AP) [4, chap. 1] if $\forall \epsilon > 0 \exists \ell_\epsilon > 0$ such that for any interval $(t_0, t_0 + \ell_\epsilon) \exists \tau_\epsilon \in (t_0, t_0 + \ell_\epsilon)$ such that $|x(t + \tau_\epsilon) - x(t)| < \epsilon \forall t \in \mathbb{R}$. The quantity τ_ϵ is called a translation number of $x(t)$ corresponding to ϵ . \square*

Each AP function $x(t)$ is associated to a Fourier series with coefficients

$$x_\lambda = \left\langle x(t) e^{-j\lambda t} \right\rangle_t \quad \lambda \in \Lambda \quad (5.1)$$

where Λ is the countable set of the possibly incommensurate frequencies [4].

Assumption 5.2. *The coefficients x_λ defined in (5.1) are such that*

$$\sum_{\lambda \in \Lambda} |x_\lambda| < \infty. \quad (5.2)$$

Proposition 5.3. *In [4] it is shown that any AP function is bounded and is the uniform limit of a sequence of trigonometric polynomial functions. Furthermore, under Assumption 5.2, we have*

$$x(t) = \sum_{\lambda \in \Lambda} x_\lambda e^{j\lambda t}. \quad (5.3)$$

Remark 5.4. *The time average of the AP function is equal to the Fourier coefficient for $\lambda = 0$, that is, $\langle x(t) \rangle_t = x_0$. Thus, the coefficient x_{λ_0} of the series expansion (5.3) of $x(t)$ is equal to the time average of $e^{-j\lambda_0 t} x(t)$ (see (5.1)). Therefore, the estimation of any Fourier coefficient x_{λ_0} can be reduced to the estimation of the time average of the AP function $e^{-j\lambda_0 t} x(t)$. Thus, the results of this section regarding the estimation of the time average hold for the estimation of any Fourier coefficient of an AP function. \square*

Lemma 5.5. *An almost-periodic function is RM [29, Theorem 6.1]. Two or more almost-periodic functions are jointly RM [29, Theorem 6.3].* \square

Lemma 5.6. *Covariance. Let be $x(t)$ given in (5.3). We have*

$$\text{cov}_t \{x(t + \tau + s_1), x(t + s_2)\} = \sum_{\lambda \neq 0} |x_\lambda|^2 e^{j\lambda\tau} e^{j\lambda(s_1 - s_2)}. \quad (5.4)$$

Proof: The result follows from a straightforward calculation. \square

Lemma 5.7. *Cumulants. Let be $x(t)$ given in (5.3). We have*

$$\text{cum}_t \{x(t + v_1), \dots, x(t + v_k)\} = \sum_{\mathbf{P}} (-1)^{p-1} (p-1)! \sum_{\lambda_1 \in \Lambda} \dots \sum_{\lambda_k \in \Lambda} \left(\prod_{i=1}^k x_{\lambda_i} e^{j\lambda_i v_i} \right) \left(\prod_{i=1}^p \delta_{\lambda_{\nu_i} \mathbf{1}} \right) \quad (5.5)$$

where \mathbf{P} is the set of distinct partitions $\{\nu_1, \dots, \nu_p\}$ of $\{1, \dots, k\}$, λ_{ν_i} is the vector whose entries are λ_ℓ with $\ell \in \nu_i$, and $\mathbf{1} \triangleq [1, \dots, 1]$. Note that in (5.5), Λ can be replaced by $\Lambda \setminus \{0\}$.

Proof: See Appendix. \square

5.2. Normalized Time Average

From Lemmas 5.6 and 5.7 we have that almost-periodic functions do not satisfy Assumptions 3.4 and 3.5. In this section, asymptotic results for normalized time average (3.8) of AP functions are established.

The estimator of the time average of the AP function (5.3) can be expressed as

$$\widehat{m}_T(t) \triangleq \frac{1}{T} \int_{t-T/2}^{t+T/2} x(u) \, du = x_0 + \sum_{\lambda \neq 0} \frac{x_\lambda}{\lambda} 2 e^{j\lambda T/2} \sin(\lambda t). \quad (5.6)$$

Lemma 5.8. *The k th-order cumulant of the normalized time average (3.8) of the almost-periodic function (5.3) under Assumption 5.2 is given by*

$$\begin{aligned} \text{cum}_t \{\widetilde{m}_T(t + \tau_1), \dots, \widetilde{m}_T(t + \tau_k)\} &= T^{k(a-1)} \sum_{\mathbf{P}} (-1)^{p-1} (p-1)! \\ &\sum_{\lambda_1 \neq 0} \dots \sum_{\lambda_k \neq 0} \left(\prod_{i=1}^k x_{\lambda_i} e^{j\lambda_i \tau_i} \right) \left(\prod_{i=1}^p \delta_{\lambda_{\nu_i} \mathbf{1}} \right) 2^k \prod_{i=1}^k \frac{\sin(\lambda_i T/2)}{\lambda_i}. \end{aligned} \quad (5.7)$$

In particular, for the covariance we have

$$\text{cov}_t \{\widetilde{m}_T(t + \tau), \widetilde{m}_T(t)\} = \text{cum}_t \{\widetilde{m}_T(t + \tau), \widetilde{m}_T(t)\} = 4T^{2(a-1)} \sum_{\lambda \neq 0} |x_\lambda|^2 e^{j\lambda\tau} \frac{\sin^2(\lambda T/2)}{\lambda^2} \quad (5.8)$$

Proof: See Appendix. \square

Assumption 5.9. *The almost-periodic function $x(t)$ with series (5.3) is such that $\lambda = 0$ is not a cluster of cycle frequencies. That is,*

$$\lambda_{\text{inf}} \triangleq \inf_{\substack{\lambda \in \Lambda \\ \lambda \neq 0}} |\lambda| > 0. \quad (5.9)$$

Lemma 5.10. *Let $x(t)$ be a non constant almost-periodic function satisfying Assumptions 5.2 and 5.9. For the normalized time average $\tilde{m}_T(t)$ with rate of convergence T^a defined in (3.8), we have that for every $k \geq 2$*

$$\lim_{T \rightarrow \infty} \text{cum}_t \{ \tilde{m}_T(t + \tau_1), \dots, \tilde{m}_T(t + \tau_k) \} = 0 \quad 0 \leq a < 1 \quad (5.10)$$

$$\lim_{T \rightarrow \infty} \text{cum}_t \{ \tilde{m}_T(t + \tau_1), \dots, \tilde{m}_T(t + \tau_k) \} \text{ does not exist} \quad a = 1 \quad (5.11)$$

$$\limsup_{T \rightarrow \infty} | \text{cum}_t \{ \tilde{m}_T(t + \tau_1), \dots, \tilde{m}_T(t + \tau_k) \} | < \infty \quad a = 1 \quad (5.12)$$

$$\lim_{T \rightarrow \infty} | \text{cum}_t \{ \tilde{m}_T(t + \tau_1), \dots, \tilde{m}_T(t + \tau_k) \} | = \infty \quad a > 1. \quad (5.13)$$

For $k = 2$ we deduce the asymptotic behavior for the covariance $\text{cov}_t \{ \tilde{m}_T(t + \tau), \tilde{m}_T(t) \}$.

Proof: See Appendix. \square

From Lemma 5.10 and Theorem 3.2, it follows that for $0 \leq a < 1$ the normalized time average $\tilde{m}_T(t)$ defined in (3.8) as $T \rightarrow \infty$ converges weakly in the FOT sense (Def. 2.8) to the distribution degenerated in zero.

From Lemma 5.10, we have that for $a = 1$ the covariance $\text{cov}_t \{ \tilde{m}_T(t + \tau), \tilde{m}_T(t) \}$ is not convergent as $T \rightarrow \infty$. However, since $| \text{cov}_t \{ \tilde{m}_T(t + \tau), \tilde{m}_T(t) \} |$ is bounded, sequences $\{ T_n \}_{n \in \mathbb{N}}$ exist such that the limit $\lim_{n \rightarrow \infty} \text{cov}_t \{ \tilde{m}_{T_n}(t + \tau), \tilde{m}_{T_n}(t) \}$ exists and is finite (and depends on the considered sequence $\{ T_n \}_{n \in \mathbb{N}}$). Analogously, we have that for $a = 1$ the cumulant $\text{cum}_t \{ \tilde{m}_T(t + \tau_1), \dots, \tilde{m}_T(t + \tau_k) \}$ is not convergent as $T \rightarrow \infty$. However, since $| \text{cum}_t \{ \tilde{m}_T(t + \tau_1), \dots, \tilde{m}_T(t + \tau_k) \} |$ is bounded (see (A.17) and following sentence), sequences $\{ T_n \}_{n \in \mathbb{N}}$ exist such that the limit $\lim_{n \rightarrow \infty} \text{cum}_t \{ \tilde{m}_{T_n}(t + \tau_1), \dots, \tilde{m}_{T_n}(t + \tau_k) \}$ exists and is finite.

In line with the above considerations, we have the following results.

Theorem 5.11. FOT Distribution of the Normalized Time Average of a Periodic Function. *Let $x(t)$ be measurable, locally integrable, and periodic with period $P > 0$. Let us consider the sequence of data-record lengths $T_n = 2nP + \rho_n$ with $\rho_n \rightarrow \rho \in [0, 2P)$ as $n \rightarrow \infty$.*

For the normalized time average (3.8) with $a = 1$ we have

$$\lim_{n \rightarrow \infty} \tilde{m}_{T_n}(t) = \lim_{n \rightarrow \infty} T_n [\hat{m}_{T_n}(t) - m] = \Psi^\rho(t) \quad (5.14)$$

where

$$\Psi^\rho(t) \triangleq \int_{t-\rho/2}^{t+\rho/2} [x(u) - m] du. \quad (5.15)$$

In addition, in the continuity points of $F_{\Psi^\rho}(\xi)$ we have

$$\lim_{n \rightarrow \infty} F_{\tilde{m}_{T_n}}(\xi) = \lim_{n \rightarrow \infty} \mu_R(\{t \in \mathbb{R} : \tilde{m}_{T_n}(t) \leq \xi\}) = \mu_R(\{t \in \mathbb{R} : \Psi^\rho(t) \leq \xi\}) \triangleq F_{\Psi^\rho}(\xi) \quad (5.16)$$

where $F_{\tilde{m}_{T_n}}(\xi)$ and $F_{\Psi^\rho}(\xi)$ are the FOT distributions of $\tilde{m}_{T_n}(t)$ and $\Psi^\rho(t)$, respectively.

Proof: See Appendix. \square

Notice that the function $\Psi^\rho(t)$ is periodic in t with period P , and periodic in ρ with period $2P$. Moreover $\Psi^0(t) = \Psi^P(t) = 0$. Thus, whenever $\rho = 0$ or $\rho = P$, the limit FOT-distribution $F_{\Psi^\rho}(\xi)$ is the Dirac measure at 0.

Example 5.12. Let be $x(t) = \cos(2\pi\nu t + \varphi)$, where $\nu > 0$ and $P = 1/\nu$. We have that

$$m = 0 \quad (5.17)$$

$$\text{cov}_t[x(t + \tau), x(t)] = \frac{1}{2} \cos(2\pi\nu\tau) \quad (5.18)$$

$$\widehat{m}_T(t) = \frac{1}{\pi\nu T} \cos(2\pi\nu t + \varphi) \sin(\pi\nu T) \quad (5.19)$$

$$\text{cov}_t[\widehat{m}_T(t + \tau), \widehat{m}_T(t)] = \frac{1}{2\pi^2\nu^2 T^2} \sin^2(\pi\nu T) \cos(2\pi\nu\tau) \quad (5.20)$$

$$\Psi^\rho(t) = \frac{1}{\pi\nu} \cos(2\pi\nu t + \varphi) \sin(\pi\nu\rho). \quad (5.21)$$

The FOT probability density of $\Psi^\rho(t)$ is

$$f_{\Psi^\rho}(\xi) = \frac{\nu}{\sqrt{\sin^2(\pi\nu\rho) - \pi^2\nu^2\xi^2}} \text{ for } |\xi| < \frac{|\sin(\pi\nu\rho)|}{\pi\nu} \text{ and 0 otherwise.} \quad (5.22)$$

Theorem 5.13. FOT Distribution of the Normalized Time Average of the Sum of Two Periodic Functions with Incommensurate Periods. Let be $x(t) = x_1(t) + x_2(t)$, where each $x_i(t)$ is locally integrable and periodic with period P_i . The periods are assumed incommensurate, that is, they are not relatively rational.

1) Let

$$T_n = 2n_{1,n}P_1 + \rho_{1,n} = 2n_{2,n}P_2 + \rho_{2,n} \quad (5.23)$$

with $n_{1,n}, n_{2,n} \rightarrow \infty$, $\rho_{1,n} \rightarrow \rho_1 \in [0, 2P_1)$, and $\rho_{2,n} \rightarrow \rho_2 \in [0, 2P_2)$ as $n \rightarrow \infty$. We have

$$\lim_{n \rightarrow \infty} \widetilde{m}_{T_n}(t) = \lim_{n \rightarrow \infty} T_n(\widehat{m}_{T_n}(t) - m) = \Psi_1^{\rho_1}(t) + \Psi_2^{\rho_2}(t) \triangleq \Psi^{\rho_1, \rho_2}(t) \quad (5.24)$$

where $\Psi_i^{\rho_i}(t)$, $i = 1, 2$ are defined according to (5.15). In addition, in the continuity points of $F_{\Psi^{\rho_1, \rho_2}}(\xi)$ we have

$$\begin{aligned} \lim_{n \rightarrow \infty} F_{\widetilde{m}_{T_n}}(\xi) &= \lim_{n \rightarrow \infty} \mu_R(\{t \in \mathbb{R} : \widetilde{m}_{T_n}(t) \leq \xi\}) \\ &= \mu_R(\{t \in \mathbb{R} : \Psi_1^{\rho_1}(t) + \Psi_2^{\rho_2}(t) \leq \xi\}) \triangleq F_{\Psi^{\rho_1, \rho_2}}(\xi) \end{aligned} \quad (5.25)$$

2) Conversely, when there is a sequence $\{T_n\}_n$ converging to infinity such that $T_n(\widehat{m}_{T_n}(t) - m)$ converges in FOT-distribution, then the limit is necessarily a FOT-distribution $F_{\Psi^{\rho_1, \rho_2}}(\xi)$ for some $0 \leq \rho_1 < 2P_1$ and $0 \leq \rho_2 < 2P_2$.

3) For each $\rho_1 \in [0, 2P_1)$ and $\rho_2 \in [0, 2P_2)$, there is a subsequence of $T_n(\widehat{m}_{T_n}(t) - m)$ which converges to $\Psi^{\rho_1}(t) + \Psi^{\rho_2}(t)$.

Proof: See Appendix. □

This theorem can be easily generalized to any almost periodic signal $x(t)$ which is the sum of a finite number of periodic functions with linearly independent periods.

Finally, note that the results in Theorems 5.11 and 5.13 are novel in the FOT approach and do not have any counterpart in the classical stochastic approach.

6. Numerical Results

In this section we present numerical results that illuminate the theory developed in Sections 4 and 5.

6.1. Bass Functions

A finite segment of the Bass function (Definition 4.11) is built with $q(t) = \text{sinc}^2(t/T_p)$, where $T_p = 4T_s$, with T_s sampling period. The Bass sequence $\{b_n\}_{n=0,\dots,N-1}$ is generated according to (4.9) by using a 3rd-order polynomial $P(n) = p_0 + p_1 n + p_2 n^2 + p_3 n^3$ with coefficients $p_0 = \sqrt{2}$, $p_1 = \sqrt{1.5}$, $p_2 = \sqrt{0.7}$, and $p_3 = \sqrt{0.15}$.

Let

$$\mu_T \triangleq \left\langle \tilde{m}_T(t) \right\rangle_t \quad \text{and} \quad \sigma_T \triangleq \left[\left\langle \left[\tilde{m}_T(t) - \mu_T \right]^2 \right\rangle_t \right]^{1/2} \quad (6.1)$$

be the FOT mean and FOT standard deviation of $\tilde{m}_T(t)$, respectively. Here $\langle \cdot \rangle_t$ denotes the average over $t \in [-Z/2, Z/2]$.

The FOT probability density function (pdf) of the normalized time-average $\tilde{m}_T(t)$ with $a = 1/2$ (see (3.8)) is evaluated for different values of data-record length $Z = NT_s$ (N even for the sake of simplicity) and T . It is assumed $T = \sqrt{N}T_s$. Therefore, in the ideal case where $Z \rightarrow \infty$ we have $T \rightarrow \infty$ with $Z/T \rightarrow \infty$.

The FOT pdf of $[\tilde{m}_T(t) - \mu_T]/\sigma_T$ (with $a = 1/2$) is computed by the kernel-based estimator [38, Sec. 2.1.8, pp. 64-65]

$$f_N^{(\text{kernel})}(\xi) = \int_{-\infty}^{+\infty} \beta_N^{-1} W\left(\frac{\xi - t}{\beta_N}\right) dF_N(t) \simeq \sum_{i=-N/2}^{N/2} W\left(\frac{\xi - \xi_i}{\beta_N}\right) \frac{1}{(N+1)\beta_N} \quad (6.2)$$

and is compared with the standard normal pdf. In (6.2), $[\xi_1, \dots, \xi_N]$ represents the sample with $\xi_i \triangleq [\tilde{m}_T(iT_s) - \mu_T]/\sigma_T$. The function $W(\xi) = e^{-\xi^2/2}/\sqrt{2\pi}$ is the kernel, and the scaling parameter β_N is chosen according to the Silverman's rule: $\beta_N = (4/3)^{1/5} N^{-1/5}$. The estimation procedure is summarized in Table 1.

The pdf $f_N^{(\text{kernel})}(\xi)$ is evaluated in $N_p = 2((Z - T)/T_s)^{1/3}$ points (Rice rule) uniformly distributed in the interval $[-5.5, 5.5]$.

The kernel-based estimator and the QQ plot of the sample data are compared with the standard normal for $N = 50 \cdot 10^3$ (Fig. 1), $N = 100 \cdot 10^3$ (Fig. 2), and $N = 500 \cdot 10^3$ (Fig. 3). The convergence of the kernel-based pdf estimator to the standard normal is evident.

compute	μ_T, σ_T by (6.1)
compute	$\xi_i = [\tilde{m}_T(iT_s) - \mu_T]/\sigma_T$ for $i = -N/2, \dots, N/2$
compute	$f_N^{(\text{kernel})}(\xi)$ by (6.2)

Table 1: FOT pdf estimation algorithm.

In order to have a quantitative measure of the convergence of the kernel-based pdf estimate to the standard normal, the Kolmogorov-Smirnov metric

$$D_N \triangleq \sup_{\xi} \left| F_N^{(\text{kernel})}(\xi) - F_{\text{normal}}(\xi) \right| \quad (6.3)$$

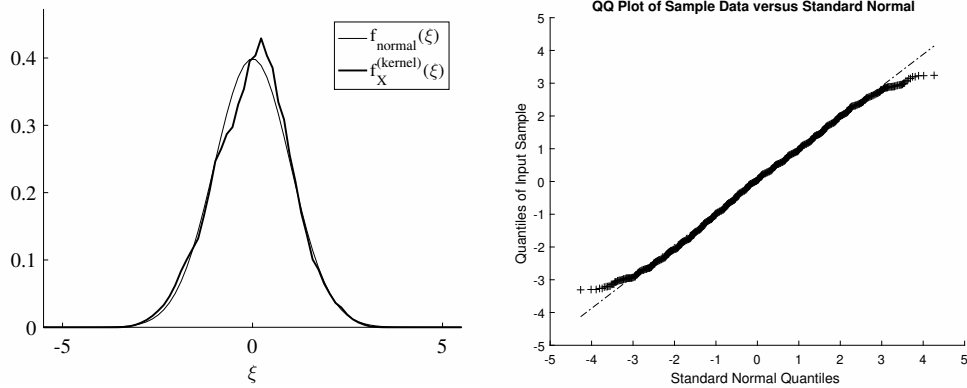


Figure 1: $N = 50 \cdot 10^3$. Left: Kernel-based estimator of the FOT pdf versus standard normal. Right: QQ plot of sample data versus standard normal.

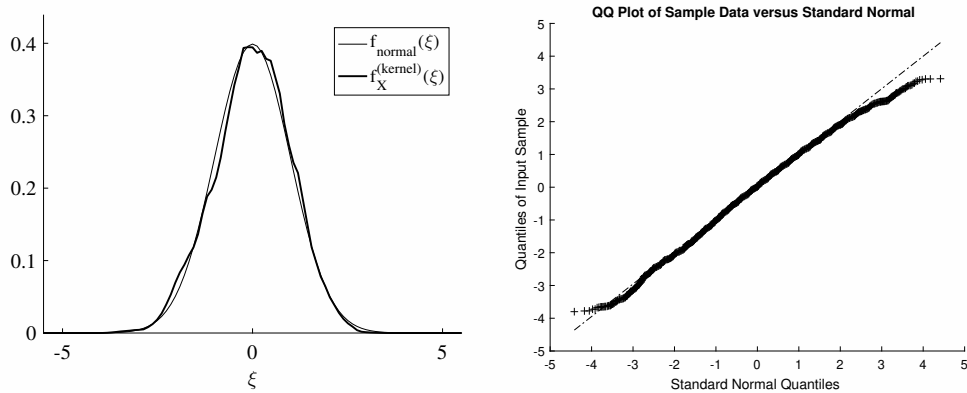


Figure 2: $N = 100 \cdot 10^3$. Left: Kernel-based estimator of the FOT pdf versus standard normal. Right: QQ plot of sample data versus standard normal.

with

$$F_N^{(\text{kernel})}(\xi) = \sum_{i=-N/2}^{N/2} \int_{-\infty}^{\xi} W\left(\frac{u - \xi_i}{\beta_N}\right) du \frac{1}{(N+1)\beta_N} \quad (6.4)$$

and the L^1 metric

$$L_N \triangleq \int_{\mathbb{R}} \left| dF_N^{(\text{kernel})}(\xi) - dF_{\text{normal}}(\xi) \right| = \int_{\mathbb{R}} \left| f_N^{(\text{kernel})}(\xi) - f_{\text{normal}}(\xi) \right| d\xi \quad (6.5)$$

are computed for $N = 50 \cdot 10^3$, $100 \cdot 10^3$ and $500 \cdot 10^3$. The results are reported in Table 2.

6.2. Almost-Periodic Functions

In order to illustrate the result of Theorem 5.13, the sum of two sinusoids with incommensurate frequencies is considered. At least one frequency must be irrational and, hence, at discrete-time, at least one sinusoid must be almost periodic. In fact, at discrete time, a sinusoid is periodic only if the frequency is a rational number. Therefore, strictly speaking, the discrete-time counterpart of (A.19) does not hold. However, a close approximate equality in the

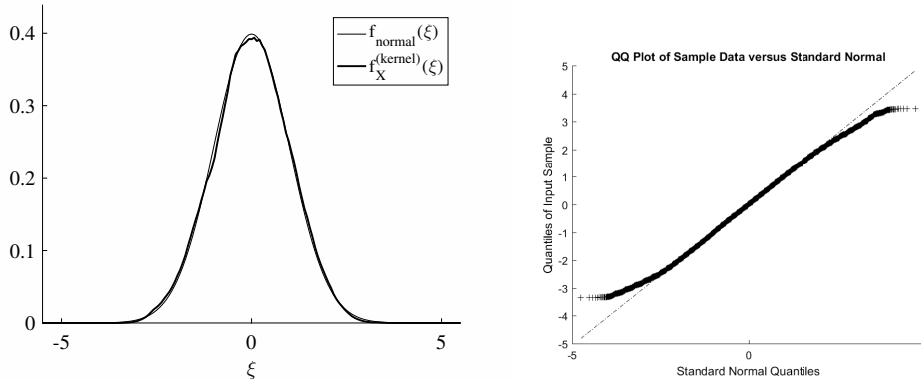


Figure 3: $N = 500 \cdot 10^3$. Left: Kernel-based estimator of the FOT pdf versus standard normal. Right: QQ plot of sample data versus standard normal.

N	D_N	L_N
$50 \cdot 10^3$	0.0208	0.0754
$100 \cdot 10^3$	0.0129	0.0587
$500 \cdot 10^3$	0.0063	0.0314

Table 2: Kolmogorov-Smirnov metric D_N and L^1 metric L_N for increasing values of N .

discrete-time counterpart of (A.19) is obtained if the discrete-time signal is obtained by sampling a continuous-time signal with sufficiently high sampling rate.

Let us consider the continuous-time signal $x(t) = \cos(2\pi f_1 t) + \cos(2\pi f_2 t)$ with $1/f_1 = P_1 = 1$ and $1/f_2 = P_2 = \pi$ sampled with period $T_s = 10^{-2}$ s. Let us assume $\rho_1 = 1/2$ and $\rho_2 = 1$.

By taking $n_{1,n} = n$ and $\rho_{1,n} = \rho_1$, the sequences T_n , $n_{2,n}$, and $\rho_{2,n}$ are given by

$$T_n = 2nP_1 + \rho_1 = 2n_{2,n}P_2 + \rho_{2,n} \quad n_{2,n} = \left\lfloor \frac{2nP_1 + \rho_1}{2P_2} \right\rfloor \quad \rho_{2,n} = 2nP_1 + \rho_1 - 2n_{2,n}P_2. \quad (6.6)$$

For each n , the integer $n_k \leq n$ is chosen such that $|\rho_{2,n_k} - \rho_2|$ is minimum. We obtain the results in Table 3.

n_k	n_{1,n_k}	ρ_{1,n_k}	n_{2,n_k}	ρ_{2,n_k}	T_{n_k} [s]
16	16	0.5	5	1.08407	32.5
195	195	0.5	62	0.942511	390.5
239	239	0.5	76	0.977917	478.5
261	261	0.5	83	0.99562	522.5
616	616	0.5	196	0.99568	1232.5
971	971	0.5	309	0.99574	1942.5

Table 3: Sequences.

The FOT probability density function (pdf) of the normalized time-average $\tilde{m}_{T_{n_k}}(t)$ with $a = 1$ (see (3.8)) is evaluated for different values of data-record length $Z = NT_s$ (N even) and T_{n_k} . It is assumed that Z grows faster than T_{n_k} as n_k increases. The kernel-based estimate

(6.2) is adopted and is compared with the theoretical pdf $f_{\text{theoretical}}(\xi)$. The two sinusoids with incommensurate frequencies are independent functions in the FOT sense [12], [26, Chap. 4], [27]. Thus,

$$f_{\text{theoretical}}(\xi) = f_{\Psi_1^{\rho_1}}(\xi) \otimes f_{\Psi_2^{\rho_2}}(\xi) \quad (6.7)$$

where \otimes denotes the convolution and $f_{\Psi_i^{\rho_i}}(\xi)$ are defined according to (5.22). In addition, for the sake of comparison, the density of $\Psi^{\rho_1, \rho_2}(t) = \Psi_1^{\rho_1}(t) + \Psi_2^{\rho_2}(t)$ is estimated with the kernel method. The kernel and the scaling parameter for the density estimation are chosen as in Section 6.1.

In Figs. 4–6, the estimated FOT pdf of $\tilde{m}_{T_{n_k}}(t)$ (thick solid line), the numerically evaluated theoretical pdf (6.7) (thin solid line), and estimated FOT pdf of $\Psi^{\rho_1, \rho_2}(t)$ (thin dotted line) are reported for $T_{n_k} = 32.5$ s, $Z = 200 T_{n_k}$ (Fig. 4), $T_{n_k} = 390.5$ s, $Z = 250 T_{n_k}$ (Fig. 5), and $T_{n_k} = 522.5$ s, $Z = 300 T_{n_k}$ (Fig. 6).

A quantitative measure of the convergence of the kernel-based pdf estimate to the theoretical pdf (6.7) when the data-record length increases along the sequence $\{T_{n_k}\}$, is obtained by computing the Kolmogorov-Smirnov metric

$$D_N \triangleq \sup_{\xi} \left| f_N^{(\text{kernel})}(\xi) - f_{\text{theoretical}}(\xi) \right| \quad (6.8)$$

and the L^1 metric

$$L_N = \int_{\mathbb{R}} \left| f_N^{(\text{kernel})}(\xi) - f_{\text{theoretical}}(\xi) \right| d\xi. \quad (6.9)$$

Results for the values of T_{n_k} considered in Figs. 4–6 are reported in Table 4.

T_{n_k} [s]	D_N	L_N
32.5	0.285	0.228
390.5	0.224	0.186
522.5	0.176	0.056

Table 4: Kolmogorov-Smirnov metric D_N and L^1 metric L_N for increasing values of T_{n_k} .

The results in Figs. 4–6 and Table 4 confirm the convergence along subsequences proved in Theorem 5.13.

In order to illustrate the convergence to different distributions depending on the values of ρ_1 and ρ_2 , an experiment with $\rho_2 = 0.4$ is carried out. In Fig. 7, the result when the approximation of the estimate to the theoretical distribution is close is reported.

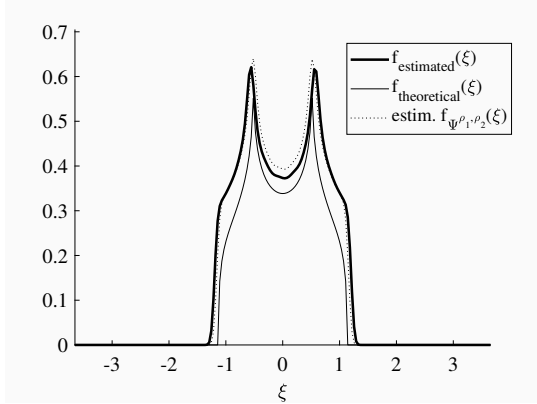


Figure 4: $T_{n_k} = 32.5$ s. Estimated FOT pdf of $\tilde{m}_{T_{n_k}}(t)$ (thick solid line), numerically evaluated theoretical pdf (6.7) (thin solid line), estimated FOT pdf of $\Psi^{\rho_1, \rho_2}(t)$ (thin dotted line).

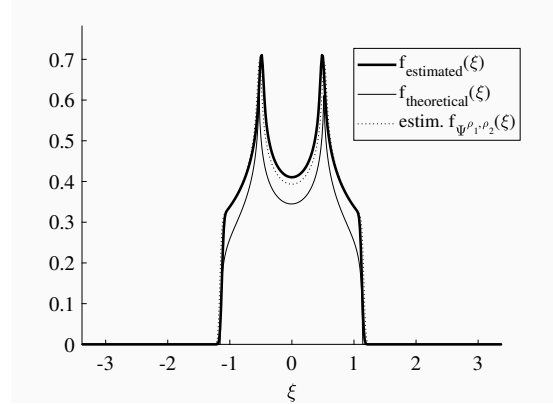


Figure 5: $T_{n_k} = 390.5$ s. Estimated FOT pdf of $\tilde{m}_{T_{n_k}}(t)$ (thick solid line), numerically evaluated theoretical pdf (6.7) (thin solid line), estimated FOT pdf of $\Psi^{\rho_1, \rho_2}(t)$ (thin dotted line).

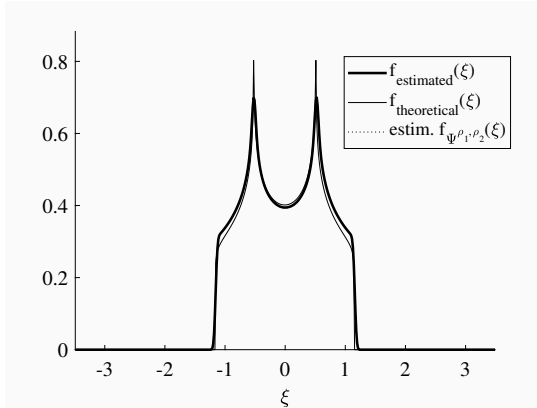


Figure 6: $T_{n_k} = 522.5$ s. Estimated FOT pdf of $\tilde{m}_{T_{n_k}}(t)$ (thick solid line), numerically evaluated theoretical pdf (6.7) (thin solid line), estimated FOT pdf of $\Psi^{\rho_1, \rho_2}(t)$ (thin dotted line).

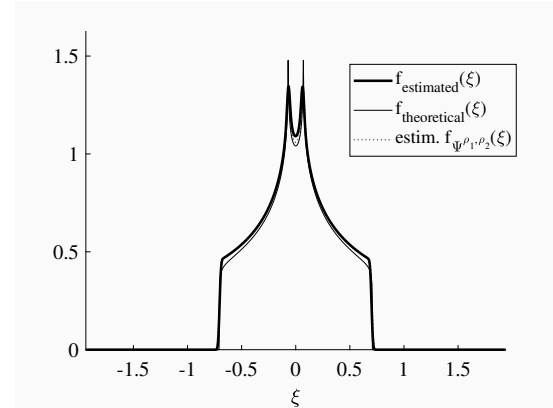


Figure 7: $T_{n_k} = 1156.5$ s, $Z = 500T_{n_k}$. Estimated FOT pdf of $\tilde{m}_{T_{n_k}}(t)$ (thick solid line), numerically evaluated theoretical pdf (6.7) (thin solid line), estimated FOT pdf of $\Psi^{\rho_1, \rho_2}(t)$ (thin dotted line).

7. Application

7.1. Test for Presence of a Non-Zero-Mean Cyclostationary Signal

In the context of cyclostationary signal analysis, the cyclic correlogram is adopted as estimator of the cyclic autocorrelation function, a key parameter adopted in a plethora of applications [35].

In the classical stochastic process framework, a properly normalized version of the cyclic correlogram can be proved to be asymptotically normal for a zero mean process under mild mixing assumptions, possibly expressed in terms of summability of cumulants [8], [11]. Thus, the result is not valid, in particular, if the process $x(t)$ contains an additive almost-periodic component.

In this section, by exploiting Theorem 3.7 and Lemma 5.10, the asymptotic normality of a properly normalized version of the cyclic correlogram is proved in the FOT framework for the case of signals possibly containing and additive almost-periodic component. Such a problem is of interest in communications when (possibly intentional) interfering tones are superposed

to the zero-mean signal of interest [35] or in machine monitoring and diagnostics [2]. Thus, a cyclostationarity detector, based on the asymptotic normality of the cyclic periodogram, originally derived in the case of zero-mean signals, is proved to be effective even in the case of strong sinewave components superimposed to the signal of interest.

Every signal can be decomposed into the sum of an almost-periodic component and a residual term not containing any finite-strength additive sinewave component. In the most general case, almost periodicity must be considered in a generalized sense [34, Sec. 6.3.1]. In the following, the decomposition is considered with some further regularity assumptions on the almost-periodic component and the residual term.

Let us consider signals for which the following decomposition holds $x(t) = x_{\text{uap}}(t) + x_r(t)$, where $x_{\text{uap}}(t)$ is a uniformly almost-periodic function with Fourier series

$$x_{\text{uap}}(t) = \sum_{\eta \in E_1} x_\eta e^{j2\pi\eta t} \quad (7.1)$$

and $x_r(t)$ is a residual term not containing any finite-strength sinewave component, that is

$$\langle x_r(t) e^{-j2\pi\eta t} \rangle_t = 0 \quad \forall \eta \in \mathbb{R}. \quad (7.2)$$

Let us assume that the Fourier series in (7.1) is absolutely convergent, the functions of any set of translates $x_r(t + s_1), \dots, x_r(t + s_k)$ are jointly RM, $x_r(t + s)$ and $x_{\text{uap}}(t)$ are jointly RM, and $x_r(t)$ is such that Assumptions 3.4 and 3.5 are satisfied and, moreover, (3.7) holds at least for $k = 4$.

The cyclic autocorrelation of $x(t)$ at cycle frequency α is [35, Eq. (2.5)]

$$R_x^\alpha(\tau) \triangleq \langle x(t + \tau) x(t) e^{-j2\pi\alpha t} \rangle_t = \sum_{\eta \in E_1} x_\eta x_{\eta - \alpha}^* e^{j2\pi\eta\tau} + R_{x_r}^\alpha(\tau). \quad (7.3)$$

With reference to the notation of Section 3, the infinite time average $m \equiv R_x^\alpha(\tau)$ is estimated by the finite-time average $\widehat{m}_T(t) \equiv R_x^{(T)}(\alpha, \tau; t)$, where

$$R_x^{(T)}(\alpha, \tau; t) \triangleq \frac{1}{T} \int_{t-T/2}^{t+T/2} x(u + \tau) x(u) e^{-j2\pi\alpha u} du \quad (7.4)$$

is the cyclic correlogram of $x(t)$. We have

$$\begin{aligned} \sqrt{T} [\widehat{m}_T(t) - m] &\equiv \sqrt{T} \left[R_x^{(T)}(\alpha, \tau; t) - R_x^\alpha(\tau) \right] \\ &= \sqrt{T} \left[\sum_{\eta_1} \sum_{\eta_2} x_{\eta_1} x_{\eta_2}^* e^{j2\pi\eta_1\tau} \frac{1}{T} \int_{t-T/2}^{t+T/2} e^{-j2\pi(\alpha - \eta_1 + \eta_2)u} du - \sum_{\eta \in E_1} x_\eta x_{\eta - \alpha}^* e^{j2\pi\eta\tau} \right] \\ &\quad + \sqrt{T} \left[\sum_{\eta} x_\eta \frac{1}{T} \int_{t-T/2}^{t+T/2} x_r(u + \tau) e^{-j2\pi(\alpha - \eta)u} du - 0 \right] \\ &\quad + \sqrt{T} \left[\sum_{\eta} x_\eta e^{j2\pi\eta\tau} \frac{1}{T} \int_{t-T/2}^{t+T/2} x_r(u) e^{-j2\pi(\alpha - \eta)u} du - 0 \right] \\ &\quad + \sqrt{T} \left[\frac{1}{T} \int_{t-T/2}^{t+T/2} x_r(u + \tau) x_r(u) e^{-j2\pi\alpha u} du - R_{x_r}^\alpha(\tau) \right] \end{aligned} \quad (7.5)$$

When $T \rightarrow \infty$, the first term approaches zero due to (5.10) (Lemma 5.10).

Since $x_r(t)$ satisfies Assumptions 3.4 and 3.5, also the real and imaginary parts of $x_r(t) e^{j2\pi\tau t}$ do. Furthermore, similarly to the case of stochastic cumulants, the k th-order FOT cumulant of the lag products $x_r(t + \tau + s_i) x_r(t + s_i)$, $i = 1, \dots, k$, can be expressed in terms of FOT cumulants of $x_r(t + s_i)$ up to order $2k$ [28], [5, Sec. 2.3], [34, pp. 83, 84]. Thus, since (3.7) with $k = 4$ and Assumptions 3.4 and 3.5 hold for $x_r(t)$, then Assumptions 3.4 and 3.5 also hold for the lag-product waveform $x_r(t + \tau) x_r(t)$. In particular, (3.7) with $k = 4$ assures the summability of the covariance of $x_r(t + \tau) x_r(t)$.

Therefore, the real and imaginary parts of second, third, and fourth term in the right-hand side of (7.5) are asymptotically jointly normal due to Theorem 3.7. Consequently, the real and imaginary parts of their sum are asymptotically jointly normal.

In conclusion, Theorem 3.7 and Lemma 5.10 allow one to establish the asymptotic joint normality of real and imaginary parts of the properly normalized cyclic periodogram even in the case of non-zero mean signal, a result that is not available in the literature in the stochastic process framework and that cannot be established in this framework.

As an example of application, let us consider the problem of testing the presence of cyclostationarity in the available data $x(t)$ by the statistical test proposed in [8] (in the stochastic process framework). The FOT counterpart of this test is based on the asymptotic joint normality of the real and imaginary parts of the function of t in (7.5). Specifically, defined the column vector $\mathbf{r} \triangleq [\text{Re}\{R_x^{(T)}(\alpha, \tau; t)\}, \text{Im}\{R_x^{(T)}(\alpha, \tau; t)\}]^T$ and its FOT covariance matrix $\mathbf{\Sigma}$, the test

$$\begin{aligned} H_0 & : R_x^\alpha(\tau) = 0 \\ H_1 & : R_x^\alpha(\tau) \neq 0 \end{aligned} \quad (7.6)$$

based on the observation of $x(u)$ for $u \in [t - T/2, t + T/2]$ reduces to

$$Q^{(T)}(t) \triangleq T \mathbf{r}^T \mathbf{\Sigma}^{-1} \mathbf{r} \underset{H_0}{\overset{H_1}{\geq}} \lambda \quad (7.7)$$

where the threshold λ is analytically determined starting from the FOT false alarm rate P_{fa} since $Q^{(T)}(t)$ has asymptotically a χ_2^2 FOT distribution $F_{\chi_2^2}(\xi)$ under H_0 . That is, we have $\lambda = F_{\chi_2^2}^{-1}(1 - P_{\text{fa}})$.

Note that test (7.7) is formally analogous to the one derived in [8] since both are obtained by straightforward calculation on Gaussian densities. However, the test statistic $Q^{(T)}(t)$ is a function of time while its analogous in [8] is a random variable and does not depend on t . In [8], the performance analysis is made in terms of probability of missed detection and probability of false alarm. In contrast, in Section 7.2, the FOT counterparts of these probabilities are defined and adopted for performance analysis in the FOT approach.

Test (7.7) is derived in the stochastic process framework under the assumption that the process $x(t)$ has summable cumulants and, in particular, is zero mean. The non-zero-mean case is addressed in [8, Sec. II-B] where two situations are considered. In the first one, the non zero mean is due to a transient added function that does not give contribution to the computation of the cyclic autocorrelation. In the second one, it is assumed that the non zero mean is consistently estimated and then subtracted from the original process $x(t)$.

The estimation of the mean is not practical if the almost-periodic component is not periodic and with many unknown frequencies. However, according to the results of this section, such estimation procedure can be avoided by performing test (7.7) directly on the available data $x(t)$.

7.2. Simulation Setup

In this section, the simulation setup in the FOT probability framework is defined and the correspondence with the analogous setup in classical stochastic framework is clarified.

For the test statistic $Q^{(T)}(t)$ defined in (7.7) the FOT probability of missed detection is defined as

$$P_{\text{md}} \triangleq \mu_{\mathbb{R}} \left(\{t \in \mathbb{R} : Q^{(T)}(t) \leq \lambda \mid \mathbf{H}_1\} \right) = \lim_{Z \rightarrow \infty} \frac{1}{Z} \int_{-Z/2}^{Z/2} \mathbf{1}_{\{Q^{(T)}(t) \leq \lambda \mid \mathbf{H}_1\}} dt. \quad (7.8)$$

It is worthwhile to underline that in the FOT framework the time variable t plays the same role played in the stochastic approach by the variable ω ranging in the sample space. Thus, the FOT probability of missed detection P_{md} is defined in terms of relative measure $\mu_{\mathbb{R}}$ and is the temporal average of the indicator of the set of values of t such that $Q^{(T)}(t) \leq \lambda$ under the hypothesis that $R_x^\alpha(\tau) \neq 0$. In contrast, in the classical stochastic approach, the probability of missed detection is defined in terms of a measure on the sample space Ω and is the ensemble average of the indicator of the event that the test statistic is less than the threshold under the hypothesis \mathbf{H}_1 .

Similarly, the FOT probability of false alarm is defined as

$$P_{\text{fa}} \triangleq \mu_{\mathbb{R}} \left(\{t \in \mathbb{R} : Q^{(T)}(t) > \lambda \mid \mathbf{H}_0\} \right) = \lim_{Z \rightarrow \infty} \frac{1}{Z} \int_{-Z/2}^{Z/2} \mathbf{1}_{\{Q^{(T)}(t) > \lambda \mid \mathbf{H}_0\}} dt. \quad (7.9)$$

The estimate of P_{md} over the finite data-record length $t \in [-Z/2, Z/2]$ is obtained as

$$P_{\text{md}}^{(Z)} = \frac{1}{Z} \int_{-Z/2}^{Z/2} \mathbf{1}_{\{Q^{(T)}(t) \leq \lambda \mid \mathbf{H}_1\}} dt \simeq \frac{1}{2M+1} \sum_{m=-M}^M \mathbf{1}_{\{Q^{(T)}(mT_s) \leq \lambda \mid \mathbf{H}_1\}} \quad (7.10a)$$

$$\simeq \frac{1}{2L+1} \sum_{\ell=-L}^L \mathbf{1}_{\{Q^{(T)}(\ell h T_s) \leq \lambda \mid \mathbf{H}_1\}}. \quad (7.10b)$$

The approximate equality in (7.10a) is due to the discrete time approximation of the continuous time average with sampling period T_s , where $2M+1$ is the largest odd integer less than Z/T_s . In (7.10a), $2M+1$ data-blocks of length $T = NT_s$ are used to compute the values $Q^{(T)}(mT_s)$ for the average. Two adjacent blocks are temporally shifted one another by T_s and overlap by $N-1$ samples. The approximate equality (7.10b) is to reduce the computational complexity of (7.10a). In fact, in (7.10b), two adjacent blocks are temporally shifted one another by hT_s and overlap by $\max\{0, N-h\}$ samples. Only $L < M$ blocks of length T are considered for computing the average, where L depends on the value of h . If $h > N$, blocks do not overlap. In such a case we have the FOT counterpart of the Monte Carlo simulations made in the classical stochastic approach.

The simulation procedure is summarized in Table 5

7.3. Numerical Results

The cyclostationary signal-of-interest (SOI) contained in $x(t)$ is obtained by filtering a PAM signal (Sec. 4.4) with stationary white binary modulating sequence, rectangular pulse with 50% duty cycle, and symbol period $T_p = 8T_s$, where $T_s = 1/f_s$ is the sampling period. The adopted filter is a one-pole system with bandwidth $f_s/8$. The SOI exhibits cyclostationarity

compute	$\mathbf{r} \triangleq [\text{Re}\{R_x^{(T)}(\alpha, \tau; t)\}, \text{Im}\{R_x^{(T)}(\alpha, \tau; t)\}]^\top$ for $t = \ell h T_s, \ell = -L, \dots, L$
estimate	Σ by moving block bootstrap [36]
compute	$Q^{(T)}(t)$ by (7.7) for $t = \ell h T_s, \ell = -L, \dots, L$
compute	$\lambda = F_2^{-1}(\chi_2^2(1 - P_{\text{fa}}))$
estimate	P_{md} by $P_{\text{md}}^{(Z)}$ given in (7.10b)

Table 5: FOT Monte Carlo simulation to estimate P_{md} .

at cycle frequencies $\alpha = k/T_p$, with k integer. Additive Gaussian noise with slowly varying power spectral level flat within the bandwidth $(-f_s/2, f_s/2)$ and two equal power sinusoidal tones are superimposed to the SOI. The frequencies of the two tones are $1/(2T_p)$ and $\sqrt{2}/(2T_p)$. Thus, their sum is almost-periodic. According to (7.3), the tone at frequency $1/(2T_p)$ gives non-zero contribution to the cyclic autocorrelation at cycle frequency $\alpha = 1/T_p$. If considered as interference, the magnitudes of the two tones, compared to the magnitude of SOI, are such that the signal-to-interference ratio (SIR) is -10 dB.

The test (7.7) is performed at $(\alpha, \tau) = (1/T_p, 0)$. $N_b = 2^8$ bits are processed to compute the cyclic correlograms. Thus, the data-record length is $N = T/T_s = N_b T_p/T_s = 2^{11}$. The covariance matrix Σ is estimated by moving block bootstrap [36] with subsamples (blocks) of length T taken when t ranges in $[-Z/2, Z/2]$. See also [13], [34, Secs. 2.6.4.1, 6.3.5].

10^5 Monte Carlo experiments are carried out to evaluate the test performance in terms of FOT probability of missed detection P_{md} as a function of the Gaussian signal-to-noise ratio (SNR) for fixed values of P_{fa} that determine the threshold λ . In the FOT context, different signal segments used for the Monte Carlo simulation are taken from a unique time series observed in a time interval $[-Z/2, Z/2]$ with $Z \gg T$. The results reported in Fig. 8, according to the theoretical results of Section 7.1, show the effectiveness of test (7.7) even in the case of strong additive almost-periodic component.

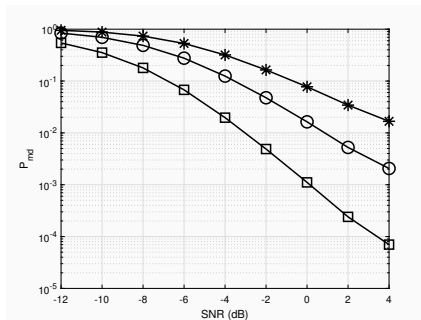


Figure 8: P_{md} as function of SNR for $(\square) P_{\text{fa}} = 10^{-2}$, $(\circ) P_{\text{fa}} = 10^{-3}$, and $(*) P_{\text{fa}} = 10^{-4}$.

8. Conclusion

Time average estimation is addressed in the fraction-of-time probability framework. For an observed signal, modeled as a single function of time, summability of temporal cumulants

is assumed. Then, a central limit theorem (Theorem 3.7) is proved for the normalized time average estimation error, with normalizing factor equal to the square root of the observation interval length. The Agnew-Kac pseudo-random functions are shown to satisfy the sufficient conditions for the CLT. These functions, however, are difficult to be implemented. Thus, as close approximations, the Bass functions and functions whose modulating sequences are obtained by a linear congruent generator are considered since they are suitable models for signals of interest in communications. The notable class of the almost-periodic functions is also considered. In such a case, it is shown that a singular distribution is obtained when the normalizing factor is the square root of the observation interval length. However, a non normal limit distribution for the normalized error of the time average estimate can be obtained when the normalizing factor equals the observation interval length. The limit distribution is shown to depend on the sequence adopted to perform the limit (Theorems 5.11 and 5.13). This result cannot be obtained in the classical stochastic approach. Several numerical examples illustrate the theoretical results. As an example of application, the asymptotic normality of the estimation error of the cyclic autocorrelation is proved in the case of non-zero mean signals and the result is exploited for cyclic detection.

Appendix A. Proofs

Proof of Theorem 3.2

Using (3.2), for every finite T and Z we have

$$\frac{1}{Z} \int_{-Z/2}^{Z/2} [\widehat{m}_T(t) - m] dt = \frac{1}{T} \int_{-T/2}^{T/2} \frac{1}{Z} \int_{-Z/2}^{Z/2} [x(u+t) - m] dt du \quad (\text{A.1})$$

where the interchange of the order of integrals is granted by the Fubini and Tonelli theorem [6, Sec. 3.4]. In fact, $x(t)$ is Lebesgue measurable and bounded and both integration intervals $[-T/2, T/2]$ and $[-Z/2, Z/2]$ are finite. In the limit as $Z \rightarrow \infty$ we have

$$\begin{aligned} \lim_{Z \rightarrow \infty} \frac{1}{Z} \int_{-Z/2}^{Z/2} [\widehat{m}_T(t) - m] dt &= \lim_{Z \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \frac{1}{Z} \int_{-Z/2}^{Z/2} [x(u+t) - m] dt du \\ &= \frac{1}{T} \int_{-T/2}^{T/2} \lim_{Z \rightarrow \infty} \frac{1}{Z} \int_{-Z/2}^{Z/2} [x(u+t) - m] dt du = 0 \end{aligned} \quad (\text{A.2})$$

where the limit in Z in the third term is zero due to (3.1). The interchange of limit and integral operations in (A.2) is allowed by the dominated convergence theorem [6, Sec. 4.3]. In fact, since $x(t)$ is bounded, for some $C > 0$ we have that $\sup_t |x(t)| < C$. Hence,

$$\left| \frac{1}{Z} \int_{-Z/2}^{Z/2} [x(u+t) - m] dt \right| \leq \frac{1}{Z} \int_{-Z/2}^{Z/2} |x(u+t) - m| dt \leq 2C$$

with the right-hand side summable in $[-T/2, T/2]$ and independent of Z .

Proof of Theorem 3.7

The asymptotic zero mean for every a is consequence of Theorem 3.2.

In the following, it is proved that for $a = 1/2$ the asymptotic covariance is finite and joint cumulants of order 3 or greater are asymptotically zero. Therefore, accounting for (2.9) (with

$x_i(t)$ replaced by $\tilde{m}_T(t + \tau_i)$), we have that, asymptotically, the logarithm of the FOT characteristic function of $\tilde{m}_T(t + \tau_i)$, $i = 1, \dots, k$, is a quadratic homogeneous polynomial in the real variables ω_i 's. That is, for $a = 1/2$, the functions $\tilde{m}_T(t + \tau_i)$ have asymptotically a joint normal FOT distribution.

For $k \geq 2$ we have the following equalities:

$$\begin{aligned}
& \text{cum}_t \{ \tilde{m}_T(t + \tau_1), \dots, \tilde{m}_T(t + \tau_k) \} \\
&= \text{cum}_t \left\{ T^{a-1} \int_{t+\tau_1-T/2}^{t+\tau_1+T/2} x(u_1) du_1, \dots, T^{a-1} \int_{t+\tau_k-T/2}^{t+\tau_k+T/2} x(u_k) du_k \right\} \\
&= T^{k(a-1)} \underbrace{\int_{-T/2}^{T/2} \dots \int_{-T/2}^{T/2}}_k \text{cum}_t \{ x(t + \tau_1 + v_1), \dots, x(t + \tau_k + v_k) \} dv_1 \dots dv_k \\
&= T^{k(a-1)} \underbrace{\int_{-T/2}^{T/2} \dots \int_{-T/2}^{T/2}}_k \text{cum}_t \{ x(t + \tau_1 + v_1 - v_k), \dots, x(t + \tau_{k-1} + v_{k-1} - v_k), x(t + \tau_k) \} dv_1 \dots dv_k \\
&= T^{k(a-1)} \int_{-T/2}^{T/2} \underbrace{\int_{-T/2-v_k}^{T/2-v_k} \dots \int_{-T/2-v_k}^{T/2-v_k}}_{k-1} \text{cum}_t \{ x(t + \tau_1 + s_1), \dots, x(t + \tau_{k-1} + s_{k-1}), x(t + \tau_k) \} ds_1 \dots ds_{k-1} dv_k
\end{aligned} \tag{A.3}$$

where in the first equality subtracting the constant value $T^a m$ to all functions of t does not modify the cumulant; in the second equality the variable changes $u_i = t + \tau_i + v_i$, $i = 1, \dots, k$ are made; in the third equality the multilinearity property of cumulants is used to interchange the cumulant $\text{cum}_t \{ \cdot \}$ and integrals $\int_{-T/2}^{T/2} (\cdot) dv_i$ operations, and all the time series in the argument of cumulant are translated of the same quantity v_k ; in the fifth equality the variable changes $v_i - v_k = s_i$, $i = 1, \dots, k - 1$, are made.

From (A.3) and accounting for Assumption 3.5 it follows that

$$\begin{aligned}
& | \text{cum}_t \{ \tilde{m}_T(t + \tau_1), \dots, \tilde{m}_T(t + \tau_k) \} | \\
&\leq T^{k(a-1)} \int_{-T/2}^{T/2} \underbrace{\int_{-T}^T \dots \int_{-T}^T}_{k-1} \left| \text{cum}_t \{ x(t + \tau_1 + s_1), \dots, x(t + \tau_{k-1} + s_{k-1}), x(t + \tau_k) \} \right| ds_1 \dots ds_{k-1} dv_k \\
&= T^{k(a-1)+1} o(T^{k/2-1})
\end{aligned} \tag{A.4}$$

as $T \rightarrow \infty$. The right-hand side of (A.4) approaches zero as $T \rightarrow \infty$ when $a = 1/2$ provided that $k \geq 3$ since k is integer.

Therefore, for $a = 1/2$, cumulants of order 3 or greater are asymptotically zero.

For $k = 2$ equality (A.3) specializes into

$$\begin{aligned}
& \text{cov}_t [\tilde{m}_T(t + \tau_1), \tilde{m}_T(t + \tau_2)] \\
&= T^{2(a-1)} \int_{-T/2}^{T/2} \int_{-T/2-v_2}^{T/2-v_2} \text{cov}_t \{ x(t + \tau_1 + s_1), x(t + \tau_2) \} ds_1 dv_2 \\
&= T^{2(a-1)} \int_{-T/2}^{T/2} \int_{\mathbb{R}} \text{cov}_t \{ x(t + \tau_1 + s_1), x(t + \tau_2) \} \text{rect} \left(\frac{s_1 + v_2}{T} \right) ds_1 dv_2
\end{aligned}$$

$$= T^{2a-1} \int_{-T}^T \text{cov}_t \left\{ x(t + \tau_1 + s_1), x(t + \tau_2) \right\} \left(1 - \frac{|s_1|}{T} \right) ds_1 \quad (\text{A.5})$$

where to obtain the third equality the Fubini and Tonelli theorem can be used since the function $s \mapsto \text{cov}_t \{x(t + \tau_1 + s), x(t + \tau_2)\}$ is bounded and Lebesgue measurable.

Under Assumption 3.4 we have that the limit of the integral in the right-hand side of (A.5) (without normalizing factor T^{2a-1}) exists and is finite. Thus, the asymptotic covariance $K(\tau_i - \tau_j)$ is finite if $a = 1/2$.

Finally, notice that

$$K(\tau_i - \tau_j) = \int_{\mathbb{R}} \text{cov}_t [x(t + \tau_i - \tau_j + s), x(t)] ds = \int_{\mathbb{R}} \text{cov}_t [x(t + \tau_i + s), x(t + \tau_j)] ds.$$

Proof of Lemma 4.3

The left-hand-side of (4.4) is the k th-order FOT cumulant (see (2.9)):

$$\text{cum}_t \{x(t + \tau_1), \dots, x(t + \tau_k)\} = (-j)^k \frac{\partial^k}{\partial \omega_1 \dots \partial \omega_k} \log \left\langle e^{j\omega_1 x(t + \tau_1)} \dots e^{j\omega_k x(t + \tau_k)} \right\rangle \Big|_{\omega=0}. \quad (\text{A.6})$$

The cumulant function in (A.6) depends on the lag differences $\tau_1 - \tau_k, \dots, \tau_{k-1} - \tau_k$. It is the time average of the k th-order FOT cumulant [33, Eq. (13)] which is defined adopting as expectation operator the almost-periodic component extraction operator (almost-cyclostationary FOT model [34, Sec. 6.3]). The pseudo random function $x(t)$ is a pulse amplitude-modulated (PAM) signal with pulse $q(t)$ and white modulating sequence having k th-order cumulant c_k . The time average of its k th-order FOT cumulant [33, Eq. (13)] is given by [39, Eq. (39)] (with $\beta = 0$ and $u_j = \tau_j - \tau_n$) which is coincident with the right-hand-side of (4.4).

Proof of Theorem 4.4

By making the variable change $t' = t - \tau_k$ into (4.4), we have

$$\text{cum}_t \{x(t + \tau_1), \dots, x(t + \tau_k)\} = \frac{c_k}{T_p} \int_{\mathbb{R}} q(t' + \tau_k) \prod_{i=1}^{k-1} q(t' + \tau_i) dt'. \quad (\text{A.7})$$

It results

$$\begin{aligned} & \text{cum}_t \{\tilde{m}_T(t + \tau_1), \dots, \tilde{m}_T(t + \tau_k)\} \\ &= T^{k(a-1)} \underbrace{\int_{-T/2}^{T/2} \dots \int_{-T/2}^{T/2}}_k \frac{c_k}{T_p} \int_{\mathbb{R}} \prod_{i=1}^k q(t + \tau_i + s_i) dt ds_1 \dots ds_k \\ &= T^{k(a-1)} \int_{-T/2}^{T/2} \underbrace{\int_{-T/2-s_k}^{T/2-s_k} \dots \int_{-T/2-s_k}^{T/2-s_k}}_{k-1} \frac{c_k}{T_p} \int_{\mathbb{R}} q(t' + \tau_k) \prod_{i=1}^{k-1} q(t' + \tau_i + u_i) dt' du_1 \dots du_{k-1} ds_k \\ &= T^{k(a-1)} \int_{-T/2}^{T/2} \underbrace{\int_{-T/2-s_k}^{T/2-s_k} \dots \int_{-T/2-s_k}^{T/2-s_k}}_{k-1} \frac{c_k}{T_p} r_q(u_1, \dots, u_{k-1}; \boldsymbol{\tau}) du_1 \dots du_{k-1} ds_k \end{aligned} \quad (\text{A.8})$$

where, in the first equality relations (3.8) and (4.4) are used, in the second equality the variable changes $u_i = s_i - s_k$, $i = 1, \dots, k-1$ and $t' = t + s_k$ are made, so that we obtain an integral not depending on s_k , and in the third equality the definition

$$r_q(u_1, \dots, u_{k-1}; \boldsymbol{\tau}) \triangleq \int_{\mathbb{R}} q(t' + \tau_k) \prod_{i=1}^{k-1} q(t' + \tau_i + u_i) dt'$$

with $\boldsymbol{\tau} \triangleq [\tau_1, \dots, \tau_k]$, is used.

For $k = 2$ and $\tau_2 = 0$, (A.8) specializes into

$$\begin{aligned} & \text{cov}_t \{ \tilde{m}_T(t + \tau_1), \tilde{m}_T(t) \} \\ &= T^{2(a-1)} \frac{c_2}{T_p} \int_{-T/2}^{T/2} \int_{\mathbb{R}} \text{rect}\left(\frac{u_1 + s_2}{T}\right) \int_{\mathbb{R}} q(t') q(t' + \tau_1 + u_1) dt' du_1 ds_2 \\ &= T^{2(a-1)} \frac{c_2}{T_p} \int_{\mathbb{R}} q(t') \int_{\mathbb{R}} q(t' + \tau_1 + u_1) \int_{-T/2}^{T/2} \text{rect}\left(\frac{u_1 + s_2}{T}\right) ds_2 du_1 dt' \\ &= T^{2a-1} \frac{c_2}{T_p} \int_{\mathbb{R}} q(t') \int_{\mathbb{R}} q(t' + \tau_1 + u_1) \left(1 - \frac{|u_1|}{T}\right) \text{rect}\left(\frac{u_1}{T}\right) du_1 dt' \end{aligned} \quad (\text{A.9})$$

where, in the second equality the order of integrals has been interchanged due to the Fubini and Tonelli theorem [6, Chap. 3]. Since $q(t) \in L^1(\mathbb{R})$, the Lebesgue dominated convergence theorem entails that for $a = 1/2$ the limit as $T \rightarrow \infty$ of the right-hand side (rhs) of equality (A.9) is finite and is equal to the rhs of (4.6).

Besides from equality (A.8) it follows that

$$\begin{aligned} & |\text{cum}_t \{ \tilde{m}_T(t + \tau_1), \dots, \tilde{m}_T(t + \tau_k) \}| \\ & \leq T^{k(a-1)} \underbrace{\int_{-T/2}^{T/2} \int_{-T/2-s_k}^{T/2-s_k} \dots \int_{-T/2-s_k}^{T/2-s_k} \frac{c_k}{T_p} |r_q(u_1, \dots, u_{k-1}; \boldsymbol{\tau})| du_1 \dots du_{k-1} ds_k}_{k-1} \\ & \leq T^{k(a-1)+1} \int_{\mathbb{R}^{k-1}} \frac{c_k}{T_p} |r_q(u_1, \dots, u_{k-1}; \boldsymbol{\tau})| du_1 \dots du_{k-1} \end{aligned} \quad (\text{A.10})$$

from which (4.5) immediately follows observing that, for every fixed $\boldsymbol{\tau}$, $r_q(u_1, \dots, u_{k-1}; \boldsymbol{\tau}) \in L^1(\mathbb{R}^{k-1})$.

Proof of Theorem 4.5

Assume that $a = 1/2$ and $k \geq 3$. From (4.5), it follows that $\lim_{T \rightarrow \infty} \text{cum}_t \{ \tilde{m}_T(t + \tau_1), \dots, \tilde{m}_T(t + \tau_k) \} = 0$. Therefore, if $a = 1/2$, asymptotically as $T \rightarrow \infty$, the covariance is finite (relation (4.6)) and all the FOT cumulants of order 3 or greater are zero. That is, the functions $\tilde{m}_T(t + \tau_1), \dots, \tilde{m}_T(t + \tau_k)$ are asymptotically jointly normal.

Proof of Theorem 4.7

The proof, reported in [1] for the continuous-time counterpart of the Agnew-Kac sequence, consists in showing that for every k and every integers r_1, \dots, r_k it results

$$\langle a_{n+m_1}^{r_1} \dots a_{n+m_k}^{r_k} \rangle_n = \langle a_{n+m_1}^{r_1} \rangle_n \dots \langle a_{n+m_k}^{r_k} \rangle_n \quad (\text{A.11})$$

which is the necessary and sufficient condition for FOT independence of k bounded sequences [27, Theorem 2], [12, Theorem 2.8].

Proof of Corollary 4.8

The k th-order cumulant of the sequence a_n , $n \in \mathbb{Z}$, can be expressed as (see (2.9) for the continuous-time counterpart definition)

$$\text{cum}_n \{a_{n+m_1}, \dots, a_{n+m_k}\} = (-j)^k \frac{\partial^k}{\partial \omega_1 \dots \partial \omega_k} \log \langle e^{j\omega_1 a_{n+m_1}} \dots e^{j\omega_k a_{n+m_k}} \rangle_n \Big|_{\omega=0} \quad (\text{A.12})$$

where $\omega = [\omega_1, \dots, \omega_k]$.

If at least two m_ℓ are different, then the corresponding shifted versions of the Agnew-Kac sequence are FOT independent (Theorem 4.7) and the time average in (A.12) factorizes into the product of two or more time averages, each containing a number strictly less than k of complex exponentials. The logarithm of the product of these time averages is equal to the sum of the logarithms of each time average. Each term of the sum does not depend on all the $\omega_1, \dots, \omega_k$ so that the k th-order derivative with respect to $\omega_1, \dots, \omega_k$ is zero.

Proof of Lemma 5.7

Using the FOT counterpart of the Leonov and Shyryaev relationship between k th-order cumulant and k th- and lower-order moments [28], we have

$$\begin{aligned} & \text{cum}_t \{x(t+v_1), \dots, x(t+v_k)\} \\ &= \sum_{\mathbf{P}} (-1)^{p-1} (p-1)! \prod_{i=1}^p \left\langle \prod_{\ell \in \nu_i} x(t+v_\ell) \right\rangle_t = \sum_{\mathbf{P}} (-1)^{p-1} (p-1)! \prod_{i=1}^p \left\langle \prod_{\ell \in \nu_i} \sum_{\lambda_\ell \in \Lambda} x_{\lambda_\ell} e^{j\lambda_\ell(t+v_\ell)} \right\rangle_t \\ &= \sum_{\mathbf{P}} (-1)^{p-1} (p-1)! \sum_{\lambda_1 \in \Lambda} \dots \sum_{\lambda_k \in \Lambda} \left(\prod_{i=1}^k x_{\lambda_i} e^{j\lambda_i v_i} \right) \prod_{i=1}^p \delta_{\lambda_{\nu_i} \mathbf{1}} \end{aligned} \quad (\text{A.13})$$

where the first summation extends over all distinct partitions $\mathbf{P} = \{\nu_1, \dots, \nu_p\}$, $p = 1, \dots, n$, of $\{1, \dots, n\}$. In the last equality, we used the fact that $\nu_i \cap \nu_j = \emptyset$ for $i \neq j$.

Finally, note that the sums over λ_ℓ can be extended to $\lambda_\ell \neq 0$ since

$$\text{cum}_t \{x(t+v_1), \dots, x(t+v_k)\} = \text{cum} \{x(t+v_1) - x_0, \dots, x(t+v_k) - x_0\}.$$

Proof of Lemma 5.8

Accounting for the multilinearity property of cumulants and (A.13), we have

$$\begin{aligned} & \text{cum}_t \{\tilde{m}_T(t+\tau_1), \dots, \tilde{m}_T(t+\tau_k)\} \\ &= T^{k(a-1)} \underbrace{\int_{-T/2}^{T/2} \dots \int_{-T/2}^{T/2}}_k \text{cum}_t \{x(t+\tau_1+s_1), \dots, x(t+\tau_k+s_k)\} ds_1 \dots ds_k \\ &= T^{k(a-1)} \sum_{\mathbf{P}} (-1)^{p-1} (p-1)! \sum_{\lambda_1 \neq 0} \dots \sum_{\lambda_k \neq 0} \left(\prod_{i=1}^k x_{\lambda_i} e^{j\lambda_i \tau_i} \right) \left(\prod_{i=1}^p \delta_{\lambda_{\nu_i} \mathbf{1}} \right) \prod_{i=1}^k \int_{-T/2}^{T/2} e^{j\lambda_i s_i} ds_i \\ &= T^{k(a-1)} \sum_{\mathbf{P}} (-1)^{p-1} (p-1)! \sum_{\lambda_1 \neq 0} \dots \sum_{\lambda_k \neq 0} \left(\prod_{i=1}^k x_{\lambda_i} e^{j\lambda_i \tau_i} \right) \left(\prod_{i=1}^p \delta_{\lambda_{\nu_i} \mathbf{1}} \right) 2^k \prod_{i=1}^k \frac{\sin(\lambda_i T/2)}{\lambda_i}. \end{aligned} \quad (\text{A.14})$$

Proof of Lemma 5.10

From (A.14) it follows that

$$\begin{aligned}
& |\text{cum}_t \{\tilde{m}_T(t + \tau_1), \dots, \tilde{m}_T(t + \tau_k)\}| \\
& \leq T^{k(a-1)} 2^k \sum_{\mathbf{P}} (-1)^{p-1} (p-1)! \sum_{\lambda_1 \neq 0} \dots \sum_{\lambda_k \neq 0} \left(\prod_{i=1}^k \frac{|x_{\lambda_i}|}{|\lambda_i|} \right) \left(\prod_{i=1}^p \delta_{\lambda_{v_i} \mathbf{1}} \right) \\
& \leq T^{k(a-1)} 2^k \sum_{\mathbf{P}} (-1)^{p-1} (p-1)! \sum_{\lambda_1 \neq 0} \frac{|x_{\lambda_1}|}{|\lambda_1|} \dots \sum_{\lambda_k \neq 0} \frac{|x_{\lambda_k}|}{|\lambda_k|}. \tag{A.15}
\end{aligned}$$

For each sum, accounting for Assumption 5.9, we have the upper bound

$$\sum_{\lambda_i \neq 0} \frac{|x_{\lambda_i}|}{|\lambda_i|} = \left[\sum_{0 < |\lambda_i| < 1} + \sum_{|\lambda_i| \geq 1} \right] \frac{|x_{\lambda_i}|}{|\lambda_i|} \leq \frac{1}{\lambda_{\text{inf}}} \sum_{0 < |\lambda_i| < 1} |x_{\lambda_i}| + \sum_{|\lambda_i| \geq 1} |x_{\lambda_i}| \tag{A.16}$$

which is finite due to Assumption 5.2. This proves the case $0 \leq a < 1$.

For $a = 1$ we have that when $T \rightarrow \infty$ the magnitude of $\text{cum} \{\tilde{m}_T(t + \tau_1), \dots, \tilde{m}_T(t + \tau_k)\}$ oscillates between 0 and

$$Q_k = 2^k \sum_{\mathbf{P}} (-1)^{p-1} (p-1)! \sum_{\lambda_1 \neq 0} \frac{|x_{\lambda_1}|}{|\lambda_1|} \dots \sum_{\lambda_k \neq 0} \frac{|x_{\lambda_k}|}{|\lambda_k|} \tag{A.17}$$

which is finite due to (A.16).

For $a > 1$ the magnitude of $\text{cum} \{\tilde{m}_T(t + \tau_1), \dots, \tilde{m}_T(t + \tau_k)\}$ approaches ∞ as $T \rightarrow \infty$.

Proof of Theorem 5.11

Let be $T_n = 2nP + \rho_n$. From the periodicity with period P of $x(t)$ it follows that

$$\tilde{m}_{T_n}(t) = \int_{t-(2nP+\rho_n)/2}^{t+(2nP+\rho_n)/2} [x(u) - m] du = \int_{t-\rho_n/2}^{t+\rho_n/2} [x(u) - m] du \triangleq \Psi^{\rho_n}(t). \tag{A.18}$$

Therefore,

$$\lim_{n \rightarrow \infty} \tilde{m}_{T_n}(t) = \int_{t-\rho/2}^{t+\rho/2} [x(u) - m] du = \Psi^\rho(t).$$

As function $(t, \rho) \mapsto \Psi^\rho(t)$ is continuous and periodic in t with period P and in ρ with period $2P$, it is uniformly continuous in \mathbb{R}^2 . Therefore the sequence of functions $\Psi^{\rho_n}(t)$ converges uniformly to $\Psi^\rho(t)$. That is $\forall t \in \mathbb{R}, \quad \forall \epsilon > 0 \exists n_\epsilon : n > n_\epsilon \Rightarrow |\Psi^{\rho_n}(t) - \Psi^\rho(t)| < \epsilon$.

Thus, for n sufficiently large we have $\Psi^\rho(t) - \epsilon < \Psi^{\rho_n}(t) < \Psi^\rho(t) + \epsilon \quad \forall t$.

Therefore, for every $\xi \in \mathbb{R}$, we have the following inclusion relationships among sets provided that n is sufficiently large

$$\{t \in \mathbb{R} : \Psi^\rho(t) \leq \xi - \epsilon\} \subseteq \{t \in \mathbb{R} : \Psi^{\rho_n}(t) \leq \xi\} \subseteq \{t \in \mathbb{R} : \Psi^\rho(t) \leq \xi + \epsilon\}$$

hence

$$\frac{1}{Z} \int_{-Z/2}^{Z/2} \mathbf{1}_{\{t \in \mathbb{R} : \Psi^\rho(t) \leq \xi - \epsilon\}} dt \leq \frac{1}{Z} \int_{-Z/2}^{Z/2} \mathbf{1}_{\{t \in \mathbb{R} : \Psi^{\rho_n}(t) \leq \xi\}} dt \leq \frac{1}{Z} \int_{-Z/2}^{Z/2} \mathbf{1}_{\{t \in \mathbb{R} : \Psi^\rho(t) \leq \xi + \epsilon\}} dt.$$

Thus, in the limit as $Z \rightarrow \infty$ we have

$$\mu_R(\{t \in \mathbb{R} : \Psi^\rho(t) \leq \xi - \epsilon\}) \leq \mu_R(\{t \in \mathbb{R} : \Psi^{\rho_n}(t) \leq \xi\}) \leq \mu_R(\{t \in \mathbb{R} : \Psi^\rho(t) \leq \xi + \epsilon\}) .$$

Since ϵ is arbitrarily small, in any point of continuity of the function $\xi \mapsto F_{\Psi^\rho}(\xi)$ we have

$$\lim_{n \rightarrow \infty} \mu_R(\{t \in \mathbb{R} : \Psi^{\rho_n}(t) \leq \xi\}) = \mu_R(\{t \in \mathbb{R} : \Psi^\rho(t) \leq \xi\}) \triangleq F_{\Psi^\rho}(\xi)$$

which is coincident with (5.16).

Proof of Theorem 5.13

1) For T_n satisfying (5.23) we have

$$\begin{aligned} \tilde{m}_{T_n}(t) &= \int_{t-(2n_{1,n}P_1+\rho_{1,n})/2}^{t+(2n_{1,n}P_1+\rho_{1,n})/2} (x_1(u) - m_1) du + \int_{t-(2n_{2,n}P_2+\rho_{2,n})/2}^{t+(2n_{2,n}P_2+\rho_{2,n})/2} (x_2(u) - m_2) du \\ &= \Psi_1^{\rho_{1,n}}(t) + \Psi_2^{\rho_{2,n}}(t) \triangleq \Psi^{\rho_{1,n},\rho_{2,n}}(t) \end{aligned} \quad (\text{A.19})$$

where the decomposition $m = m_1 + m_2$ is not unique.

The functions $(t, \rho_i) \mapsto \Psi_i^{\rho_i}(t)$, $i = 1, 2$ are uniformly continuous in \mathbb{R}^2 . Hence if $\rho_{1,n} \rightarrow \rho_1$ and $\rho_{2,n} \rightarrow \rho_2$, then $\lim_{n \rightarrow \infty} \sup_t |\Psi_i^{\rho_{i,n}}(t) - \Psi_i^{\rho_i}(t)| = 0$ and

$$\begin{aligned} \lim_{n \rightarrow \infty} \sup_t |\Psi^{\rho_{1,n},\rho_{2,n}}(t) - \Psi^{\rho_1,\rho_2}(t)| &= \lim_{n \rightarrow \infty} \sup_t |\Psi_1^{\rho_{1,n}}(t) - \Psi_1^{\rho_1}(t) + \Psi_2^{\rho_{2,n}}(t) - \Psi_2^{\rho_2}(t)| \\ &\leq \lim_{n \rightarrow \infty} \sup_t |\Psi_1^{\rho_{1,n}}(t) - \Psi_1^{\rho_1}(t)| + \lim_{n \rightarrow \infty} \sup_t |\Psi_2^{\rho_{2,n}}(t) - \Psi_2^{\rho_2}(t)| = 0 . \end{aligned} \quad (\text{A.20})$$

By reasoning as in the proof of Theorem 5.11, for ϵ arbitrarily small and n sufficiently large we have that $\Psi^{\rho_1,\rho_2}(t) - \epsilon < \Psi^{\rho_{1,n},\rho_{2,n}}(t) < \Psi^{\rho_1,\rho_2}(t) + \epsilon \quad \forall t$. Consequently, in any point of continuity of the function $\xi \mapsto F_{\Psi^{\rho_1,\rho_2}}(\xi)$ we have

$$\lim_{n \rightarrow \infty} \mu_R(\{t \in \mathbb{R} : \Psi^{\rho_{1,n},\rho_{2,n}}(t) \leq \xi\}) = \mu_R(\{t \in \mathbb{R} : \Psi^{\rho_1,\rho_2}(t) \leq \xi\}) .$$

2) From the above considerations we see that the FOT-distribution of $T_n(\widehat{m}_{T_n} - m)$ converges. Moreover, for any n there exists $\rho_{1,n} \in [0, 2P_1)$ and $\rho_{2,n} \in [0, 2P_2)$ such that $T_n = 2n_{1,n}P_1 + \rho_{1,n} = 2n_{2,n}P_2 + \rho_{2,n}$.

Since the sequences $\rho_{1,n}$ and $\rho_{2,n}$ are bounded, from the Bolzano-Weierstrass theorem there exists a subsequence $\{T_{n_k}\}$ such that $\rho_{1,n_k} \rightarrow \rho_1$ and $\rho_{2,n_k} \rightarrow \rho_2$ as $k \rightarrow \infty$ for some $\rho_1 \in [0, 2P_1]$ and $\rho_2 \in [0, 2P_2]$. Then from item 1) we have

$$\lim_{n \rightarrow \infty} T_{n_k}(\widehat{m}_{T_{n_k}}(t) - m) = \Psi_1^{\rho_1}(t) + \Psi_2^{\rho_2}(t) .$$

Recall that $\Psi_i^{\rho_i}(t)$ is periodic in ρ_i with period $2P_i$, for $i = 1, 2$. Then by uniqueness of the limit in the FOT-distribution, we deduce that the limit FOT-distribution of $T_n(\widehat{m}_{T_n} - m)$ is equal to $F_{\Psi^{\rho_1,\rho_2}}(\xi)$.

3) Let now $0 \leq \rho_1 < 2P_1$ and $0 \leq \rho_2 < 2P_2$ be fixed. Let us build a subsequence of $T(\widehat{m}_T(t) - m)$ that converges to $\Psi^{\rho_1, \rho_2}(t)$.

For $n \in \mathbb{N}$, let $n_{1,n} = n$, $\rho_{1,n} = \rho_1$, $n_{2,n} \triangleq \lfloor \frac{2nP_1 + \rho_1}{2P_2} \rfloor$, and $\rho_{2,n} \triangleq 2nP_1 + \rho_1 - 2n_{2,n}P_2$.

Thus $2nP_1 + \rho_1 = 2n_{2,n}P_2 + \rho_{2,n}$ with $0 \leq \rho_{2,n} < 2P_2$. Besides, since P_1 and P_2 are linearly independent (incommensurate), the real number $\frac{nP_1}{P_2}$ is irrational. Then from the Weyl Criterion [43], the sequence $\{\frac{\rho_{2,n}}{2P_2}\}_{n \in \mathbb{N}}$ is equidistributed in the interval $[0, 1)$ and the sequence $\{\rho_{2,n}\}_{n \in \mathbb{N}}$ is equidistributed in $[0, 2P_2)$. As a consequence, for each $k > 0$ there exists $n_k > n_{k-1}$, with $n_0 = 0$, such that $|\rho_{2,n_k} - \rho_2| \leq k^{-1}$, so $\rho_{2,n_k} \rightarrow \rho_2$ as $k \rightarrow \infty$. Then, accounting for item1), the sequence $\{T_k\}_{k \in \mathbb{N}}$ defined by $T_{n_k} \triangleq 2n_kP_1 + \rho_1 = 2n_{2,n_k}P_2 + \rho_{2,n_k}$ is such that

$$\lim_{n \rightarrow \infty} T_{n_k}(\widehat{m}_{T_{n_k}}(t) - m) = \Psi_1^{\rho_1}(t) + \Psi_2^{\rho_2}(t) = \Psi^{\rho_1, \rho_2}(t).$$

References

- [1] R. P. Agnew and M. Kac, "Translated functions and statistical independence," *Bulletin of the American Mathematical Society*, vol. 47, no. 2, pp. 148–154, 1941.
- [2] J. Antoni, "Cyclostationarity by examples," *Mechanical Systems and Signal Processing*, vol. 23, no. 4, pp. 987–1036, May 2009.
- [3] J. Bass, *Cours de Mathématique*. Paris: Masson & Cie, 1971, vol. III.
- [4] A. S. Besicovitch, *Almost Periodic Functions*. London: Cambridge University Press, 1932, (New York: Dover Publications, Inc., 1954).
- [5] D. R. Brillinger, *Time Series. Data Analysis and Theory*. San Francisco, CA: Holden Day, Inc., 1981, SIAM 2001.
- [6] D. C. Champeney, *A Handbook of Fourier Theorems*. New York: Cambridge University Press, 1990.
- [7] J. W. Dalle Molle, M. J. Hinich, and D. J. Morrice, "Higher-order cumulant spectral-based statistical tests of pseudo-random variate generators," in *Proceedings of the 24th Conference on Winter Simulation*, ser. WSC '92. New York, NY, USA: ACM, 1992, pp. 618–625.
- [8] A. V. Dandawaté and G. B. Giannakis, "Statistical tests for presence of cyclostationarity," *IEEE Transactions on Signal Processing*, vol. 42, no. 9, pp. 2355–2369, 1994.
- [9] R. B. Davies, "Inference from non-ergodic time series," *Advances in Applied Probability*, vol. 11, no. 2, pp. 261–262, 1979.
- [10] D. Dehay and H. L. Hurd, "Empirical determination of the frequencies of an almost periodic time series," *Journal of Time Series Analysis*, vol. 34, no. 2, pp. 262–279, 2013.
- [11] D. Dehay and J. Leśkow, "Functional limit theory for the spectral covariance estimator," *Journal Appl. Prob.*, vol. 33, pp. 1077–1092, 1996.
- [12] D. Dehay, J. Leśkow, and A. Napolitano, "Central limit theorem in the functional approach," *IEEE Transactions on Signal Processing*, vol. 61, no. 16, pp. 4025–4037, August 15 2013.
- [13] D. Dehay, A. Dudek, and J. Leskow, "Subsampling for continuous-time almost periodically correlated processes," *Journal of Statistical Planning and Inference*, vol. 150, pp. 142–158, 2014.
- [14] J. L. Doob, *Stochastic Processes*. New York: John Wiley & Sons, Inc., 1953.

- [15] T. S. Durrani, A. R. Leyman, and J. J. Soraghan, “Reducing third order cumulant spread in computer generated random numbers,” in *IEE Colloquium on Mathematical Aspects of Digital Signal Processing*, February 1994, pp. 7/1–7/6.
- [16] V. Filipovic, N. Nedic, and V. Stojanovic, “Robust identification of pneumatic servo actuators in the real situations,” *Forschung im Ingenieurwesen*, vol. 75, no. 4, pp. 183–196, December 2011.
- [17] W. A. Gardner, *Introduction to Random Processes with Applications to Signals and Systems*. New York: Macmillan, 1985, (1990, 2nd Edn., McGraw-Hill, New York).
- [18] —, *Statistical Spectral Analysis: A Nonprobabilistic Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1987.
- [19] W. A. Gardner and W. A. Brown, “Fraction-of-time probability for time-series that exhibit cyclostationarity,” *Signal Processing*, vol. 23, pp. 273–292, June 1991.
- [20] W. A. Gardner and C. M. Spooner, “The cumulant theory of cyclostationary time-series. Part I: Foundation,” *IEEE Transactions on Signal Processing*, vol. 42, pp. 3387–3408, December 1994.
- [21] P. Halmos, “The foundations of probability,” *The American Mathematical Monthly*, vol. 51, no. 9, pp. 493–510, November 1944.
- [22] —, *Lectures on Ergodic Theory*. Mathematical Society of Japan, 1956, and American Mathematical Society (AMS) Chelsea Publishing, New York, 2006.
- [23] T. E. Hull and A. R. Dobell, “Random number generators,” *SIAM Review*, vol. 4, no. 3, pp. 230–254, 1962.
- [24] H. L. Hurd and T. Koski, “The Wold isomorphism for cyclostationary sequences,” *Signal Processing*, vol. 84, pp. 813–824, May 2004.
- [25] D. Isaacson, “Ergodicity, geometric ergodicity and strong ergodicity,” *Advances in Applied Probability*, vol. 12, no. 2, pp. 300–300, 1980.
- [26] M. Kac, *Statistical Independence in Probability, Analysis and Number Theory*. USA: The Mathematical Association of America, 1959.
- [27] M. Kac and H. Steinhaus, “Sur les fonctions indépendantes (IV) (Intervalle infini),” *Studia Mathematica*, vol. 7, pp. 1–15, 1938.
- [28] V. P. Leonov and A. N. Shiryaev, “On a method of calculation of semi-invariants,” *Theory of Probability and its Applications*, vol. 4, no. 3, pp. 319–329, 1959.
- [29] J. Leśkow and A. Napolitano, “Foundations of the functional approach for signal analysis,” *Signal Processing*, vol. 86, no. 12, pp. 3796–3825, December 2006.
- [30] —, “Non-relatively measurable functions for secure-communications signal design,” *Signal Processing*, vol. 87, no. 11, pp. 2765–2780, November 2007.
- [31] P. M. Mäkilä, J. R. Partington, and T. Norlander, “Bounded power signal spaces for robust control and modeling,” *SIAM Journal on Control and Optimization*, vol. 37, no. 1, pp. 92–117, 1998.
- [32] P. Mäkilä, “LTI approximation of nonlinear systems via signal distribution theory,” *Automatica*, vol. 42, no. 6, pp. 917 – 928, 2006.
- [33] A. Napolitano, “Cyclic higher-order statistics: input/output relations for discrete- and continuous-time MIMO linear almost-periodically time-variant systems,” *Signal Processing*, vol. 42, no. 2, pp. 147–166, March 1995.

- [34] —, *Generalizations of Cyclostationary Signal Processing: Spectral Analysis and Applications*. John Wiley & Sons Ltd - IEEE Press, 2012.
- [35] —, “Cyclostationarity: New trends and applications,” *Signal Processing*, vol. 120, pp. 385–408, March 2016.
- [36] D. N. Politis, “Computer intensive methods in statistical analysis,” *IEEE Signal Processing Magazine*, vol. 15, no. 1, pp. 39–55, January 1998.
- [37] M. Rosenblatt, *Stationary Sequences and Random Fields*. Boston, MA: Birkhäuser, 1985.
- [38] R. J. Serfling, *Approximation Theorems of Mathematical Statistics*. USA: John Wiley & Sons, Inc., 1980.
- [39] C. M. Spooner and W. A. Gardner, “The cumulant theory of cyclostationary time-series. Part II: Development and applications,” *IEEE Transactions on Signal Processing*, vol. 42, pp. 3409–3429, December 1994.
- [40] V. Stojanovic and V. Filipovic, “Adaptive input design for identification of output error model with constrained output,” *Circuits, Systems, and Signal Processing*, vol. 33, no. 1, pp. 97–113, January 2014.
- [41] V. Stojanovic and N. Nedic, “Robust identification of OE model with constrained output using optimal input design,” *Journal of the Franklin Institute*, vol. 353, no. 2, pp. 576–593, 2016.
- [42] V. Stojanovic, N. Nedic, D. Prsic, and L. Dubonjic, “Optimal experiment design for identification of ARX models with constrained output in non-Gaussian noise,” *Applied Mathematical Modelling*, vol. 40, no. 13, pp. 6676–6689, 2016.
- [43] H. Weyl, “Über die Gleichverteilung von Zahlen mod. Eins,” *Mathematische Annalen*, vol. 77, pp. 313–352, 1916.
- [44] N. Wiener, “Generalized harmonic analysis,” *Acta Mathematica*, vol. 55, pp. 117–258, 1930.