Cyclostationary Processes With Evolving Periods and Amplitudes

Soumya Das ^D and Marc G. Genton ^D

Abstract—Wide-sense cyclostationary processes are an important class of non-stationary processes that have a periodic structure in their first- and second-order moments. This article extends the notion of cyclostationarity (in the wide sense) to processes where the mean and covariance functions might depart from strict periodicities and constant amplitudes. Specifically, we propose a novel and flexible class of processes that allows periods and amplitudes of the mean and covariance functions to evolve and, therefore, accommodates a much larger class of processes than the classical cyclostationary processes. Thereafter, we investigate its properties, provide methodologies for statistical inference, and illustrate the presented methods using synthetic signals and a physical signal, from the heavens, of the magnitudes of the light emitted from the variable star R Hydrae.

Index Terms—Amplitude, cyclostationarity, non-stationarity, periodic correlation, signal processing, time deformation.

I. INTRODUCTION

T HE brightnesses of variable stars, as perceived from Earth, are signals that change slowly over time. While some are nearly periodic, others may not be quite so. Their amplitudes or periods or even both may evolve over time, although the basic underlying structures of the signals remain unaltered; see [1]–[4]. For example, the period and amplitude of the signal of R Hydrae, which is a Mira-type variable star in the constellation Hydra, declined slowly in the period 1900–1950; see [5] and [6].

Another interesting aspect of variable stars, apropos of our current study, is that the light emitted from these stars may have periodically varying temporal correlations. For example, consider the following two possible phenomena: 1) A variable star has an orbiting companion that eclipses it in T days. For ease of explanation, we consider T = 365 days or one Earth year. Now, the emitted light measured on 15 January may be statistically different from that measured on 1 May. However, it can be reasonably presumed that the emitted light measured on 1 May of different years is more or less identical and closely correlated. That is to say, the light emitted from such a variable star may display periodicity in its correlation structure. 2) A Cepheid variable star periodically swells and shrinks, pulsating

Manuscript received May 30, 2020; revised November 20, 2020; accepted January 28, 2021. Date of publication February 4, 2021; date of current version March 12, 2021. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Vincent Y. F. Tan. This work was supported by the King Abdullah University of Science and Technology (KAUST). (*Corresponding author: Soumya Das.*)

The authors are with the Statistics Program, King Abdullah University of Science and Technology, Thuwal 23955, Saudi Arabia (e-mail: soumya.das@kaust.edu.sa; marc.genton@kaust.edu.sa).

Digital Object Identifier 10.1109/TSP.2021.3057268

in brightness. Therefore, for a similar reason as above, it may also show periodicity in its correlation structure. Now, because the amplitudes and the periods of some variable stars fluctuate over time, their correlation structures may also exhibit similar departures. Here, we seek an approach to model such phenomena, as discussed in the following paragraph.

We know that the signals observed from a physical system (over time) that consists of both periodicity and randomness may exhibit cyclostationarity; see [7]–[9]. However, as stated above, some celestial phenomena require the amplitudes or periods or both to change over time, unlike the classical cyclostationary processes. The inadequate modeling of such phenomena motivates us to extend the concept of cyclostationarity to processes that have first- and second-order moments departed from strict periodicities and constant amplitudes. We call such a process an evolving period and amplitude cyclostationary (EPACS) process.

In the literature, however, there are several instances, where the authors relaxed the strict properties of cyclostationarity in many different ways to fit into different real-life scenarios. For example, [10] extended the notion of cyclostationary processes to cyclo-non-stationary processes, that still goes through periodic modulations whereas being nonstationary on a long term basis. To describe such signals in rotating machinery with non-uniform rotation speed, [11] showed that the optimal way is to consider angle and time jointly instead of independently, and proposed an angle-time covariance function in this regard; see also [12] and [13]. The angle-time covariance function, with its angle parameter an appropriate function of time, can also be viewed as our covariance model when restricted to a constant amplitude. Amongst a few other extensions of the cyclic covariance functions, [14] studied a specific case of that proposed in this article by considering a constant amplitude; see also [15]. In contrast, [16] considered only an amplitude modulated cyclostationary process. Also, while our analysis deals with all the cycle frequencies, [55] investigated elaborately a single cycle exploitation of a similar process.

Since EPACS processes are more flexible in nature than cyclostationary processes, they exhaust a larger class of processes. For example, processes with discrete scale invariances belong to this class; see Subsection II-B of this article and [17] for an introduction to such processes. Therefore, in addition to modeling certain celestial phenomena, the EPACS processes can also be used to model a wide variety of real-life scenarios relating to, for example, critical phenomena [18], climatology, environmetrics, and the growth and genealogy of populations [17].

1053-587X © 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information. After introducing the EPACS processes, the article presents their different properties and builds a necessary theory to conduct statistical inference. We provide semi-parametric methodologies to estimate the mean and covariance functions of an EPACS process and to test relevant hypotheses. To illustrate the proposed methodologies, we provide one synthetic and one physical signal example. For the physical example, we consider the signal of the light emitted from the variable star R Hydrae in the period 1900–1950 from [6]. In that paper, the authors used the signal to fit a regression model where the mean had an evolving period and amplitude structure but the errors were considered uncorrelated. We will later show in the application that R Hydrae has also an evolving period and amplitude correlation structure that was not considered in their paper.

The remainder of this article is structured as follows. Section II formally introduces the EPACS processes and presents some immediate properties. Section III builds methodologies for statistical inference, which are illustrated using synthetic signals in Section IV and using a physical signal in Section V. Section VI summarizes the findings of this article and discusses its significance, limitations, and possible future extensions.

II. EPACS PROCESSES

A. Preliminaries and Definitions

Let $\{X_t, t \in \mathbb{D}\}\$ be a univariate, second-order (i.e., finite second moment), continuous-time stochastic process with mean and covariance functions m(t) and R(s,t), respectively. Then, the process $\{X_t, t \in \mathbb{D}\}\$ is called wide-sense (or second-order) cyclostationary or periodically correlated with period T if, for every $s, t \in \mathbb{D}$,

and

$$m(t) = \mathbf{E}(X_t) = m(t+T)$$

$$R(s,t) = \operatorname{cov}(X_s, X_t) = R(s+T, t+T)$$

holds, where $\mathbb{D} \subseteq \mathbb{R}$ and T is a fixed positive real number [19]–[21]; see also [22]–[25] for an introduction to higher-order cyclostationarity. Hereafter, by "cyclostationary" processes, we will only refer to "wide-sense (or second-order) cyclostationary" processes.

We generalize the notion of the cyclostationary processes, as in Definition 1 below, to processes with time-varying periods and amplitudes. Thereafter, Definition 2 defines jointly evolving period and amplitude cyclostationary processes.

Definition 1: A second-order stochastic process $\{X_t, t \in \mathbb{D}\}$ is called an evolving period and amplitude cyclostationary (abbreviated as EPACS) process if, for every $s, t \in \mathbb{D}$,

$$m(t) = l(t)m_0(g(t)) \text{ and } R(s,t) = l(s)l(t)R_0(g(s),g(t))$$
(1)

holds, where $\mathbb{D} \subseteq \mathbb{R}$, $m_0(t)$ and $R_0(s, t)$ are bounded periodic functions with unit periods, that is,

$$m_0(t) = m_0(t+1)$$
 and $R_0(s,t) = R_0(s+1,t+1), \quad s,t \in \mathbb{D}$

l(t) is a positive smooth function, and g(t) is a continuously differentiable, strictly increasing function such that g'(t) > 0

for all $t \in \mathbb{D}$. We call l(t) the amplitude modulation, g(t) the phase function, and g'(t) the instantaneous frequency at time point t of the process $\{X_t, t \in \mathbb{D}\}$.

Definition 2: The second-order stochastic processes $\{X_t^j, j = 1, ..., N, t \in \mathbb{D}\}$ are called jointly EPACS if, for every j, k = 1, ..., N and $s, t \in \mathbb{D}$,

$$m_k(t) = \mathbf{E}\left(X_t^k\right) = l_k(t)m_0^k\left(g_k(t)\right)$$

and

$$R_{jk}(s,t) = \operatorname{cov}\left(X_{s}^{j}, X_{t}^{k}\right) = l_{j}(s)l_{k}(t)R_{0}^{jk}\left(g_{j}(s), g_{k}(t)\right)$$

holds, where $m_0^k(t), k = 1, ..., N$, and $R_0^{jk}(s,t), j, k = 1, ..., N$, are bounded periodic functions with unit periods for all $s, t \in \mathbb{D}$, $l_k(t), k = 1, ..., N$, are positive smooth functions, and $g_k(t), k = 1, ..., N$, are continuously differentiable, strictly increasing functions such that $g'_k(t) > 0, k = 1, ..., N$, for all $t \in \mathbb{D}$.

Now, it is clear from Definition 1 above that the EPACS processes have the potential to explain a wider range of real-life phenomena than the cyclostationary processes. We also note that the assumption there of m_0 and R_0 being unit-periodic does not restrict generality as, for any *T*-periodic functions m_0 and R_0 satisfying (1), we can always replace g(t) by g(t)/T, $m_0(t)$ by $m_0(Tt)$, and $R_0(s,t)$ by $R_0(Ts,Tt)$ to satisfy our assumption of unit periodicity. Therefore, we avoid the difficulties of detecting the period *T*, unlike the case of usual cyclostationarity. Also, in this article, we limit ourselves to EPACS processes with non-zero means.

Remark 1: (Non-identifiability issue) While (1) generalizes the notion of the cyclostationary processes, it also admits the drawback of non-identifiability. It is well known that the representation of a given function might not be unique; even a proper periodic function can be expressed as a function with time-varying amplitude and phase. For example, $\cos(2\pi t)$ can also be represented as $(1 + a(t)) \cos(2\pi (t + f(t)))$ for appropriate a(t) and f(t); see [26].

To cope with this identifiability issue, we need to determine the amplitude using a simple parametric function that changes slower than the lengths of periods over time. We will suggest parametric models for the amplitude modulation later in Section III and impose l(t) = 1 at the initial time point to incorporate the initial amplitude in $m_0(t)$ and not to duplicate in l(t); see [6].

Next, we present two immediate observations through the following propositions. In particular, Proposition 2 below states a key connection between the EPACS and the cyclostationary processes. For Proposition 2, it is required to establish that the function $R_0(s,t)$ is a valid covariance function; Proposition 1 ascertains this requirement.

Proposition 1: Given that R(s,t) in (1) is a covariance function, the unit-periodic function $R_0(s,t)$ is also a covariance function.

Proof: See Appendix A.

Proposition 2: The process $\{X_t, t \in \mathbb{R}\}$ is EPACS satisfying (1) if and only if the time-deformed process $\{Y_t : Y_t = X_{q^{-1}(t)}/l(g^{-1}(t)), t \in \mathbb{R}\}$ is cyclostationary with period one, that is,

$$E(Y_t) = m_0(t)$$
 and $cov(Y_s, Y_t) = R_0(s, t)$

holds.

Proof: The proof is trivial, as we recall that g(t) is an injective function in \mathbb{R} and l(t) is positive in its entire domain.

B. $(H, \exp(1))$ -DSI Processes as an Example of EPACS Processes

A real-valued stochastic process $\{X_t, t > 0\}$ is said to have a discrete scale invariance with scaling exponent H and scale λ (abbreviated as (H, λ) -DSI) if

$$\{X_{\lambda t}, t > 0\} \stackrel{\mathcal{L}}{=} \{\lambda^H X_t, t > 0\}$$

holds, where $\stackrel{\mathcal{L}}{=}$ denotes "equality of the finite dimensional distributions"; see [17]. This class of processes can be viewed as a specific case of the self-similar processes [27]–[34].

In light of the above definition, we note that if $\{X_t, t > 0\}$ is $(H, \exp(1))$ -DSI, then

$$\{Y_t = c_0 \exp(-tH) X_{\exp(t)}, t \in \mathbb{R}\}\$$

is a unit-periodic cyclostationary process and, conversely, if $\{Y_t, t \in \mathbb{R}\}$ is unit-periodic cyclostationary then

$$\{X_t = \frac{t^H}{c_0} Y_{\ln t}, t > 0\}$$

is a $(H, \exp(1))$ -DSI process, where c_0 is any non-zero real number; see [17].

This further implies that if the process $\{Y_t, t \in \mathbb{R}\}$ has mean and covariance functions $m_Y(t)$ and $R_Y(s, t)$, respectively, then the $(H, \exp(1))$ -DSI process $\{X_t, t > 0\}$ can also be viewed as an EPACS process with mean and covariance functions

$$m_X(t) = \frac{t^H}{c_0} m_Y(\ln t)$$
 and $R_X(s,t) = \frac{s^H t^H}{c_0^2} R_Y(\ln s, \ln t)$

respectively. That is, the different functions of (1) can be identified as $l(t) = \frac{t^H}{c_0}$, $g(t) = \ln t$, $m_0(t) = m_Y(t)$, and $R_0(s, t) = R_Y(s, t)$. The constant c_0 , however, should be chosen cautiously to avoid the non-identifiability issue, as mentioned in Remark 1. For example, we may choose $c_0 = t_0^H$, where $t_0 > 0$ is the initial time point.

C. Properties of EPACS Processes

Below, we present some properties of the EPACS process (1), which can also be used as models to generate EPACS processes. While stating the following properties, we assume that the amplitude modulation and the phase function are appropriately defined as in Definition 1. Proofs of the following properties are provided in Appendix B.

Property 2.3.1: Let $\{Y_t, t \in \mathbb{R}\}$ be a random periodic signal satisfying $Y_t \stackrel{\mathcal{L}}{=} Y_{t+1}$ for all $t \in \mathbb{R}$. Then, the stochastic process $\{X_t = l(t)Y_{g(t)}, t \in \mathbb{R}\}$ is EPACS with amplitude modulation l(t) and phase function g(t).

Property 2.3.2: Uncorrelated EPACS processes with the same amplitude and phase functions are closed under addition.

Property 2.3.3: EPACS processes are closed under multiplication by a scaled and shifted version of a unit-periodic deterministic signal $\{h_t, t \in \mathbb{R}\}, \{f_t = e(t)h_{g(t)}, t \in \mathbb{R}\}$, with a phase function identical to that of the initial process; here, e(t) represents the corresponding amplitude function.

Property 2.3.4: Let $\{X_t^j, j = 1, ..., N, t \in \mathbb{R}\}$ be jointly EPACS processes with identical amplitude and phase functions l(t) and g(t), respectively, and consider deterministic signals $\{f_t^j = e(t)h_{g(t)}^j, j = 1, ..., N, t \in \mathbb{R}\}$, where $h_t^j = h_{t+1}^j$. Then the process $Y_t = \sum_{j=1}^N f_t^j X_t^j$ is EPACS. Property 2.3.5: The scaled and shifted process $\{Y_t = X_t^j, Y_t^j\}$

Property 2.3.5: The scaled and shifted process $\{Y_t = a(t)X_{f(t)}, t \in \mathbb{R}\}$ of an EPACS process $\{X_t, t \in \mathbb{R}\}$ satisfying (1) is also EPACS. In particular, if $a(t) = (l(g^{-1}(t)))^{-1}$ and $f(t) = g^{-1}(t)$, then the process $\{Y_t, t \in \mathbb{R}\}$ becomes cyclostationary with period one.

D. Representation of EPACS Processes

We know that the cyclostationary signals are produced from the mixing of stationarity and periodicity; see [9]. Since EPACS processes are a simple extension of cyclostationary processes, we expect to observe a similar behavior upto changing amplitude and phase functions. The two propositions in this subsection lead to an infinite-dimensional stationary representation of the EPACS process (1).

However, before presenting the propositions, we recall that a unitary operator U on a Hilbert space \mathcal{H} satisfies $\langle Ux, Uy \rangle = \langle x, y \rangle$ for $x, y \in \mathcal{H}$. For the unitary operators considered hereafter, the time domain \mathcal{H}_Y is considered as

$$\mathcal{H}_Y = \overline{\operatorname{sp}}\{Y_t : Y_{g(t)} = X_t/l(t), t \in \mathbb{R}\},\$$

where the closure is in the mean-square sense.

Proposition 3: A second-order process $\{X_t, t \in \mathbb{R}\}$ is EPACS if and only if there exists a unitary operator V and a unit-periodic signal $\{P_t, t \in \mathbb{R}\}$ taking values in \mathcal{H}_Y such that, for every t,

$$X_t = l(t)V^{g(t)}P_{q(t)} \tag{2}$$

holds, where l(t) and g(t) are the amplitude and phase functions, respectively, of the EPACS process $\{X_t, t \in \mathbb{R}\}$.

Proof: See Appendix C.

Proposition 4: A second-order process $\{X_t, t \in \mathbb{R}\}$ is EPACS satisfying (1) if and only if there exists an infinitedimensional stationary signal $\mathbf{Z}_t = \{Z_t^j\}$ such that

$$X_t = \sum_{j=-\infty}^{\infty} l(t) Z_{g(t)}^j \exp\left(i2\pi j g(t)\right). \tag{3}$$

Proof: See Appendix D.

III. SEMI-PARAMETRIC INFERENCE FOR EPACS PROCESSES

We discuss methodologies to estimate the mean and covariance functions of an EPACS process and to test relevant hypotheses. To that end, we shall adopt a semi-parametric approach. This is because a complete parametric model for the mean and covariance functions will limit the novel flexibility of such processes and, on the other side, a complete non-parametric method may make the estimation and the testing impossible if no additional assumption is made. To be specific, we determine the functions l(t) and g(t) parametrically and the functions $m_0(t)$ and $R_0(s,t)$ non-parametrically.

A. Models for Phase and Amplitude Functions

We note that if g(t) is the phase function of a unit-periodic function h(t) at time point t, then g'(t) can be seen as the instantaneous frequency and 1/g'(t) as the instantaneous period of h(t) at that time point t (provided g'(t) is sufficiently small for all t belonging to the observation time interval). This statement follows from the observation that, for small u > 0,

$$h(g(t+u)) = h(g(t) + g'(t)u + o(u)),$$

which uses Taylor series expansion of $g(\cdot)$ around t. That is,

$$h(g(t+1/g'(t))) \approx h(g(t)+1) = h(g(t));$$

see [35], [36].

Now, to make the model easily interpretable, it is helpful to consider the phase function g in such a way that the time-varying period 1/g'(t) is fairly simple. One possible choice of such a phase function could be

$$g(t|\boldsymbol{\alpha}) = \alpha_2^{-1} \log \left(1 + \alpha_1^{-1} \alpha_2 t\right), \alpha_1 > 0, \alpha_2 \in \mathbb{R},$$

for all $t > -\alpha_1/\alpha_2$ if $\alpha_2 > 0$ and $t < -\alpha_1/\alpha_2$ if $\alpha_2 < 0$ (the conditions on the time domain are required to ensure that the argument of the logarithm and $g'(t|\alpha)$ are positive). This implies that $1/g'(t|\alpha) = \alpha_1 + \alpha_2 t$. That is, the period of the corresponding EPACS signal is α_1 at time t = 0 and it changes linearly with respect to time. Similarly, the choice

$$g(t|\boldsymbol{\alpha}) = \alpha_1^{-1}t + \alpha_2 t^2, \alpha_1 > 0, \alpha_2 \in \mathbb{R},$$

for all $t > -(2\alpha_1\alpha_2)^{-1}$ if $\alpha_2 > 0$ and $t < -(2\alpha_1\alpha_2)^{-1}$ if $\alpha_2 < 0$, suggests a reciprocal linear period (with the initial period α_1) and can easily be extended to higher-order polynomials. Another noteworthy choice of g is

$$g(t|\boldsymbol{\alpha}) = (\alpha_1 \alpha_2)^{-1} \{1 - \exp(-\alpha_2 t)\}, \alpha_1 > 0, \alpha_2 \in \mathbb{R}$$

for all $t \in \mathbb{R}$. In this case, the time-varying period, $1/g'(t|\alpha) = \alpha_1 \exp(\alpha_2 t)$, (with the initial period α_1) has a simple exponential structure. Finally, we note that all these three models boil down to a simple linear model, $g(t|\alpha) = t/\alpha_1$ for all $t \in \mathbb{R}$, as $\alpha_2 \to 0$.

In most realistic scenarios, when a signal is observed over a time interval [0, n] at an approximately constant rate, only the initial period (i.e., α_1 for the models considered above) would be considered fixed with respect to n; the other parameters (i.e., α_2 for the models considered above) are generally assumed decreasing to 0 with n. More specifically, α_2 should decrease in such a way that $n|\alpha_2| < \infty$. This ensures $g(t|\alpha)$ does not change by an order of magnitude over the observation time interval; see [6].

The choices for l can be similar to the above. But, to avoid identifiability issues, as discussed in Remark 1, we impose l(t) = 1 at time t = 0. Taking this into account, one possible choice of the amplitude modulation could be a simple linear model:

$$l(t|\beta) = 1 + \beta t, \beta \in \mathbb{R},$$

for all $t > -1/\beta$ if $\beta > 0$ and $t < -1/\beta$ if $\beta < 0$. The constraints here on the time domain are required to ensure $l(t|\beta) > 0$.

Another typical choice could be an exponential model:

$$l(t|\boldsymbol{\beta}) = \exp\left(\beta_1 t + \dots + \beta_q t^q\right), \beta_1, \dots, \beta_q \in \mathbb{R},$$

which takes, by construction, positive values throughout the real line.

For the models of the amplitude modulation also, the parameters should be considered in such a way that $\sup_n \sup_j n^j |\beta_j| < \infty$; see [6].

B. Estimation

The parametric model assumptions of the amplitude and phase functions boil down the estimation problem to estimating the parameters α and β and the functions $m_0(t)$ and $R_0(s, t)$. While we might use non-parametric methods, such as kernels, directly to model these functions, to begin with we focus on estimating the parameters α and β using an iterative algorithm, as presented below. This is because, once the parameters are estimated, we can always obtain the time deformation of the underlying process using Proposition 2 to make it cyclostationary and therefore use the standard techniques to estimate the functions $m_0(t)$ and $R_0(s,t)$. This observation also provides a basis for the algorithm, which essentially finds an α and a β , which together make the deformed process the "most" cyclostationary. Similar methods of making the deformed process the most cyclostationary were also adopted in [35], [36]. However, the algorithms presented there consider only a single cycle frequency, where the underlying cyclostationary or almost-cyclostationary signal exhibits strong cyclostationarity. In contrast, the method proposed here accounts for all the cycle frequencies of the underlying cyclostationary signal. This can be seen as an advantage of our algorithm while, at the same time, being aware of its drawback of regarding the measurements of weak harmonics in spite of being, generally, more noisy.

Consider a trajectory of an EPACS process $\{(t, x_t), t \in \mathbb{D}\}$; we write the trajectory in this way to emphasize that the data point x_t is observed at time point t. Here, \mathbb{D} denotes a set of discrete time points in \mathbb{R} , not necessarily regularly spaced, at which the data points are collected. The algorithm to estimate α and β then goes as follows:

- Step 1) Fix $\alpha = \alpha_0$ and $\beta = \beta_0$.
- Step 2) Conduct a time deformation $r_t = g(t|\alpha_0)$ and consider $y_t = x_t/l(t|\beta_0)$ of the signal $\{(t, x_t), t \in \mathbb{D}\}$ to obtain $\{(r_t, y_t), t \in \mathbb{D}\}$. Let $r_{\min} = \min_{t \in \mathbb{D}} r_t$ and $r_{\max} = \max_{t \in \mathbb{D}} r_t$.
- Step 3) Smooth $\{(r_t, y_t), t \in \mathbb{D}\}$, i.e., $\{(r_t, y_t), r_t \in [r_{\min}, r_{\max}]\}$, using an interpolating cubic spline and denote the produced curve by f(t).
- Step 4) Divide the time frame $[0, \lfloor r_{\max} \rfloor]$ into $\lfloor r_{\max} \rfloor$ disjoint intervals, each of length one unit, where $\lfloor r_{\max} \rfloor$ denotes the greatest integer not greater than r_{\max} .

- Step 5) In the *k*th interval of the time frame, predict the values of f(t) at points $\{r_t \lfloor r_t \rfloor + k 1, t \in \mathbb{D}\}$ for $k = 1, \ldots, \lfloor r_{\max} \rfloor$.
- Step 6) Estimate $m_0(r_t)$, for all $t \in \mathbb{D}$, by $\widehat{m}_0^{(0)}(r_t | \boldsymbol{\alpha}_0, \boldsymbol{\beta}_0) = \frac{1}{\lfloor r_{\max} \rfloor} \sum_{k=1}^{\lfloor r_{\max} \rfloor} f(r_t \lfloor r_t \rfloor + k 1).$

Step 7) Calculate the residual sum of squares (RSS),

$$\operatorname{RSS} = \sum_{t \in \mathbb{D}} \{ x_t - l(t|\boldsymbol{\beta}_0) \widehat{m}_0^{(0)}(r_t|\boldsymbol{\alpha}_0, \boldsymbol{\beta}_0) \}$$

Step 8) Change the values of α and β in Step 1 and redo Steps 2–7 until the minimum RSS is found. Denote the corresponding parameters that minimize the RSS by $\hat{\alpha}$ and $\hat{\beta}$.

We minimize the RSS in Step 8 using a differential evolution algorithm, which can be found readily implemented in the DEOptim package [37] in the statistical software R [38]. The differential evolution algorithm requires an objective function and the lower and upper bounds of the concerned parameter space as inputs. In our case, the objective function is the RSS as defined in Step 7 above. The lower and upper bounds of the parameter space, from our experience, can be any feasible finite vectors. For example, if we consider the amplitude and phase functions,

and

 $l(t|\beta) = 1 + \beta t, \beta \in \mathbb{R},$

$$g(t|\boldsymbol{\alpha}) = \alpha_2^{-1} \log \left(1 + \alpha_1^{-1} \alpha_2 t\right), \alpha_1 > 0, \alpha_2 \in \mathbb{R},$$

respectively, then the DEoptim is capable of searching the target vector (that minimizes the RSS) from a domain $[(-M, 0, -M)^T, (M, M, M)^T]$ of $(\beta, \alpha_1, \alpha_2)^T$, where *M* could be sufficiently large such as 10^3 . However, as we also mentioned in Subsection III-A, the parameters other than α_1 could be considered decreasing to 0 with the sample size *n* in most realistic scenarios (i.e., of the form k/n, e.g., $\beta = 1/n$ and $\alpha_2 = 20/n$). We also remark that, in certain scenarios, the algorithm may result in boundary values. This problem, however, can easily be resolved by running the DEoptim algorithm a few times (say, 10 times) and taking the best outcome (i.e., the vector corresponding to the minimum RSS amongst those 10 replicates).

After the estimates of the parameters are found, we may transform the trajectory to make it cyclostationary and use standard techniques to estimate the mean and covariance functions $m_0(t)$ and $R_0(s,t)$, respectively. However, for the mean function $m_0(t)$ we can consider the estimator corresponding to the final parameters $\hat{\alpha}$ and $\hat{\beta}$ obtained in Step 6. For the estimation of $R_0(s,t)$, readers are referred to [39, p. 147] where the author discusses a "direct" and an "indirect" method to estimate the covariance function of any cyclostationary process; see also [9], [40]–[42].

Because the estimates of the parameters α and β are produced through an iterative algorithm with many intermediate steps, they are not explicit. Therefore, although the algorithm is easy to implement, a rigorous proof of theoretical properties, such as consistency, is difficult. Nevertheless, using a simulation study as presented in Section IV, we show that the estimators of the parameters are indeed consistent; that is, as the sample size increases, the estimates tend to the true values of the parameters.

On a lighter note, if one estimates the mean function $m_0(t)$ using a kernel method for given α and β and then minimizes the mean integrated squared error to obtain the parameters, the consistency may be proved using a similar way to that given in [6]. However, for simplicity, we consider the above algorithm, which produces satisfactory numerical results; see Section IV. Also, we remark that, in a closely related context of time-warped cyclostationary or almost-cyclostationary processes, [35], [36], [43] presented methods based on maximizing an appropriately modified version of the cyclic correlogram to estimate the timewarping function, which can be viewed as an analogous to our phase function, g(t); see also [44].

C. Hypotheses Tests

To present our approach, we consider a trajectory of an EPACS process $\{(t, x_t), t \in \mathbb{D}\}$ with amplitude and phase functions, for ease of understanding and brevity,

$$l(t|\beta) = 1 + \beta t, \beta \in \mathbb{R}$$

and

$$g(t|\boldsymbol{\alpha}) = \alpha_2^{-1} \log \left(1 + \alpha_1^{-1} \alpha_2 t\right), \alpha_1 > 0, \alpha_2 \in \mathbb{R}$$

respectively. However, the discussion, as in the sequel, can be modified accordingly for other choices of $l(t|\beta)$ and $g(t|\alpha)$. We recall, from Subsection III-A, that the above choice of phase function results in a linear (time-varying) period,

$$1/g'(t|\boldsymbol{\alpha}) = \alpha_1 + \alpha_2 t.$$

Given the above choices of amplitude and phase functions, hypotheses that might be of interest are to test for

1) constant amplitude (i.e., $H_0: \beta = 0$),

2) constant period (i.e., $H_0: \alpha_2 = 0$), and

3) cyclostationarity (i.e., $H_0: \beta = 0$ and $\alpha_2 = 0$).

We recall from the previous subsection that an EPACS process for given α and β can be deformed to a cyclostationary process. Following the same idea, to test a hypothesis (say, H_0), we deform the underlying process under H_0 and check if the resulting signal is cyclostationary. A similar idea was implemented in [33] to test for the self-similarity index H of a self-similar process where the authors exploited the link between self-similar and stationary processes given by Lamperti [45] and transformed the signal under H_0 to test for stationarity.

To circumvent any possible doubt, we remark that the deformation of an EPACS process with incorrect choices of parameters must not lead to a cyclostationary process (this can be clearly seen from the construction), unlike the scenario in [33] for testing the self-similarity index. Regarding the test for cyclostationarity, one may choose amongst the many available test procedures; see, for example, [46]–[50]. Here, we use a slightly modified version of [49], where the authors evaluated the diagonal exceedance percentages of the sample squared coherences from a nominal threshold; we recall that the sample squared coherence of a signal $\{z_t, t = 1, ..., N\}$ is defined as [47]

$$|\mathcal{C}(u,v,M)|^{2} = \frac{\left|\sum_{m=0}^{M-1} d_{N}(\omega_{u+m})\overline{d_{N}(\omega_{v+m})}\right|^{2}}{\sum_{m=0}^{M-1} |d_{N}(\omega_{u+m})|^{2} \sum_{m=0}^{M-1} |d_{N}(\omega_{v+m})|^{2}}$$

where $d_N(\omega_k) = \frac{1}{\sqrt{N}} \sum_{t=1}^N z_t \exp(-i\omega_k t)$ is the discrete Fourier transform of $\{z_t, t = 1, ..., N\}$, M is the smoothing parameter, and $\omega_k = 2\pi k/N$ for k = 0, ..., N - 1.

To explain the algorithm, we test for a constant amplitude. Therefore, our hypothesis of interest can be formulated as

$$H_0: \beta = 0$$
 vs. $H_1: \beta \neq 0$

and tested using the following steps:

- Step 1) Obtain the estimates of β , α_1 , and α_2 using the method described in the previous subsection. Denote the corresponding estimates by $\hat{\beta}$, $\hat{\alpha}_1$, and $\hat{\alpha}_2$, respectively.
- Step 2) Deform the signal using the estimates to have $\{(r_t = g(t|\widehat{\alpha}_1, \widehat{\alpha}_2), y_t = x_t/l(t, \widehat{\beta})), t \in \mathbb{D}\}$ and, under H_0 , $\{(r_t = g(t|\widehat{\alpha}_1, \widehat{\alpha}_2), y_{t,H_0} = x_t/l(t, 0)), t \in \mathbb{D}\}.$
- Step 3) Smooth both the signals and interpolate at equidistant time points over the interval $[0, r_{\text{max}}]$ each at ϵ unit distances (the distance between points should be chosen in a way to ensure that we have enough points for further proceedings). Therefore, the lengths of these two deformed and regularly spaced signals become $l = |r_{\text{max}}/\epsilon| + 1$.
- Step 4) Compute the sample squared coherence, $|\mathcal{C}(u, v, M)|^2$, for all $h = |u - v| \in \{1, \dots, |l/2| -$ 1}, and calculate the ratio of the number of sample squared coherences that exceed a pre-defined threshold along each diagonal line $h \ge 1$ to the number of points to be evaluated on that diagonal line [49] for both the deformed signals, $\{(r_t, y_t), t \in \mathbb{D}\}\$ and $\{(r_t, y_{t,H_0}), t \in \mathbb{D}\}.$ Denote the corresponding diagonal exceedance ratios by $\{p_h, h = 1, \dots, |l/2| - 1\}$ and $\{q_h, h = 1, \dots, |l/2| - 1\}$, respectively.
- Step 5) Reject the null hypothesis if the Euclidean distance (hereafter denoted by $D(\cdot, \cdot)$) between the two vectors of ratios $\boldsymbol{p} = (p_1, \dots, p_{\lfloor l/2 \rfloor - 1})^T$ and $\boldsymbol{q} = (q_1, \dots, q_{\lfloor l/2 \rfloor - 1})^T$ is greater than a pre-specified positive scalar. That is, if

$$\mathbf{D}^2(\boldsymbol{p},\boldsymbol{q}) = \|\boldsymbol{p} - \boldsymbol{q}\|_2 > \delta.$$

A choice of δ will later be illustrated in Section IV. We remark that Step 1 of this algorithm depends on the effectiveness of the previous algorithm presented in Subsection III-B. In Section IV, nevertheless, we establish the efficacy of that algorithm numerically.

IV. SYNTHETIC EXAMPLE

To illustrate the theories developed in Section III, we present a detailed synthetic example as follows.

A. General Settings

The unit-periodic function m_0 was chosen to be sinusoidal, that is, $m_0(t) = \sin(2\pi t)$ for $t \in \mathbb{R}$. To understand the effect of the underlying correlation structure on the inference procedure, three different models for R_0 were considered, namely,

- 1) $R_0(s,t) = \exp\{-|s-t|\},\$
- 2) $R_0(s,t) = (1+|s-t|) \exp\{-|s-t|\},$ and
- 3) $R_0(s,t) = \min\{s,t\},\$

for $(s,t) \in [0,1]^2$, and extended to the two-dimensional (real) space by periodicity. In other words, we considered the above correlation structures with their arguments 's' and 't' replaced by 's modulo 1' and 't modulo 1,' respectively. We remark that the first two models of R_0 are stationary autocovariance functions, which can be found in many different contexts [51], while the third model can be viewed as the autocovariance function of the Wiener process within the time interval [0, 1]. For the amplitude and phase functions, we chose

$$l(t|\beta) = 1 + \beta t, \beta \in \mathbb{R}$$

and

$$g(t|\boldsymbol{\alpha}) = \alpha_2^{-1} \log \left(1 + \alpha_1^{-1} \alpha_2 t\right), \alpha_1 > 0, \alpha_2 \in \mathbb{R},$$

respectively. We took *n* regularly spaced time points over the interval [0, n] with n = 100, 400, 700, and 1000, and the corresponding signals were generated from a Gaussian distribution, $\mathcal{N}_n(\boldsymbol{m}, \boldsymbol{R})$, where $\boldsymbol{m} = (m(1), \dots, m(n))^T$ and $\boldsymbol{R} = (R(s, t))_{s,t=1}^n$. Finally, the parameters of interest were chosen to be

$$\beta = n^{-1}, \alpha_1 = 20, \text{ and } \alpha_2 = 20n^{-1}$$

B. Relative Biases and Mean Squared Errors of the Estimators

The relative biases and mean squared errors (MSEs), which were calculated over 200 simulations, of the parameters are presented in Table I; the relative bias and MSE of an estimator $\hat{\theta}$ of some representative parameter θ are calculated as

$$\operatorname{bias}(\widehat{\theta}) = \frac{\frac{1}{200}\sum_{j=1}^{200}\widehat{\theta}_j - \theta}{\theta} \text{ and } \operatorname{MSE}(\widehat{\theta}) = \frac{\frac{1}{200}\sum_{j=1}^{200}\left(\widehat{\theta}_j - \theta\right)^2}{\theta^2},$$

respectively, where $\hat{\theta}_j$ is the estimate of θ in the *j*th simulation for j = 1, ..., 200.

To study the numerical properties, we used the statistical software R [38]. The smoothing of the deformed signal (see Step 3 of the algorithms in Subsection III-B and III-C) was carried out using the splinefun function from the stats package with method fmm [52], as this method interpolates at unknown intermediate time points significantly well. To minimize the RSS, we used the differential evolution algorithm DEoptim from the DEoptim package [37].

We take away from this numerical study, as depicted in Table I, that the algorithm works quite effectively to estimate the parameters and improves itself as n increases. In other words, the proposed estimator has the much-anticipated quality of being both unbiased and consistent.

1585

TABLE I Relative Biases and MSEs of the Estimators of the Parameters Using the Algorithm Provided in Subsection III-B

n	$bias(\widehat{\beta})$	$bias(\widehat{\alpha}_1)$	$bias(\widehat{\alpha}_2)$	$MSE(\hat{\beta})$	$MSE(\hat{\alpha}_1)$	$MSE(\widehat{\alpha}_2)$
$B_0(s, t) = \exp(- s-t)$						
100^{-100}	-1.59×10^{-2}	-2.04×10^{-3}	8.60×10^{-3}	5.38×10^{-3}	4.38×10^{-5}	3.07×10^{-4}
400	4.04×10^{-3}	5.65×10^{-5}	2.50×10^{-4}	1.18×10^{-3}	2.78×10^{-7}	1.42×10^{-6}
700	3.94×10^{-3}	2.26×10^{-4}	-6.76×10^{-4}	9.16×10^{-4}	1.64×10^{-7}	1.23×10^{-6}
1000	9.14×10^{-4}	4.67×10^{-5}	-1.90×10^{-4}	4.52×10^{-4}	2.10×10^{-8}	1.37×10^{-7}
$R_0(s,t) = (1 + s - t) \exp(- s - t)$						
100	-9.16×10^{-3}	-4.18×10^{-3}	1.43×10^{-2}	8.49×10^{-4}	3.05×10^{-5}	3.50×10^{-4}
400	1.10×10^{-3}	6.90×10^{-5}	3.06×10^{-4}	2.27×10^{-5}	2.87×10^{-8}	2.40×10^{-7}
700	1.49×10^{-3}	3.63×10^{-4}	-1.07×10^{-3}	1.04×10^{-4}	2.27×10^{-7}	1.90×10^{-7}
1000	1.19×10^{-3}	5.50×10^{-5}	-2.25×10^{-4}	3.45×10^{-5}	6.45×10^{-9}	9.05×10^{-8}
$R_0(s,t) = \min\{s,t\}$						
100	1.17×10^{-2}	-2.72×10^{-3}	1.12×10^{-2}	3.60×10^{-3}	5.28×10^{-5}	4.25×10^{-4}
400	9.24×10^{-3}	1.10×10^{-4}	2.42×10^{-4}	6.11×10^{-4}	2.09×10^{-7}	1.15×10^{-6}
700	-4.45×10^{-3}	3.44×10^{-4}	-9.91×10^{-4}	5.63×10^{-4}	2.26×10^{-7}	1.78×10^{-6}
1000	-4.24×10^{-3}	4.62×10^{-5}	-1.94×10^{-4}	5.97×10^{-4}	1.99×10^{-8}	1.15×10^{-7}

C. Transforming the Signal to Cyclostationary

Now that the parameters were estimated, one might want to deform a synthetic signal into a cyclostationary signal to conduct further analysis. We checked that a deformed EPACS signal was indeed cyclostationary using one representative synthetic signal $\{(t, x_t), t \in \mathbb{D}\}$ of size n = 400 and with covariance function $R_0(s, t) = \exp\{-|s - t|\}$; here, \mathbb{D} represents the set of regularly spaced time points over the interval [0, n] of size n. The estimates of the parameters, $\beta = 1/400 = 0.0025$, $\alpha_1 = 20$, and $\alpha_2 = 20/400 = 0.05$, were obtained to be

$$\hat{\beta} = 0.00\,244, \hat{\alpha}_1 = 19.99, \text{ and } \hat{\alpha}_2 = 0.05,$$

respectively. Therefore, the signal was deformed to have $\{(r_t = g(t|\hat{\alpha}_1, \hat{\alpha}_2), y_t = x_t/l(t|\hat{\beta})), t \in \mathbb{D}\}\)$. The top panel of Fig. 1 displays the synthetic signal and the bottom panel plots the resulting cyclostationary signal after deformation. From the plot, it is clear – at least visually – that the amplitude and the period of the deformed signal became constant, unlike the initially simulated signal.

To detect the length of the period, we then plotted the periodogram (see Figure 2) of the deformed signal interpolated at regularly spaced time points. We interpolated at every $\epsilon = 0.05$ unit distant time points using an interpolating cubic spline over the interval $[0, r_{\text{max}}]$ to have $\{z_t, t = 1, \dots, N\}$ with $r_{\text{max}} =$ 13.86 and N = 278. The periodogram plot displays a significant peak at frequency ≈ 0.051 , implying the period to be 19.

Finally, to show that the deformed signal, $\{z_t, t = 1, ..., N\}$, was indeed cyclostationary with period T = 19, we plotted the sample squared coherence; see Figure 3. In a sample squared coherence plot, lines parallel to the diagonal (each at h = N/Tunit distances) indicate that the signal is cyclostationary with period T. Figure 3 displays light black lines parallel to the main diagonal periodically at each 14th unit distance. Therefore,





Fig. 1. Plot of a representative synthetic signal $\{(t, x_t), t \in \mathbb{D}\}$ (top) and the corresponding deformed signal $\{(r_t, y_t), t \in \mathbb{D}\}$ (bottom).

we conclude that the deformed signal was cyclostationary with period $278/14 \approx 19$.

D. Model Fitting

As the previous subsection concludes that the deformed signal was cyclostationary, one might further fit an appropriate periodic moving average autoregressive (PARMA) model (see [9]



Fig. 2. Periodogram plot of the deformed (regularly spaced) series $\{z_t, t = 1, \dots, N\}$.



Fig. 3. Squared coherence plot of the deformed (regularly spaced) series $\{z_t, t = 1, \ldots, N\}$.

and [53]) to the signal. We will not go into much detail about fitting the model, but for the sake of completeness, we fitted a periodic autoregressive model of order two (PAR(2)) to the demeaned signal, $\tilde{z}_t = z_t - m_z(t)$, with $m_z(t) = \frac{1}{L} \sum_{j=0}^{L-1} z_{t+jT}$, $t = 1, \ldots, T$, and L denoting the number of cycles present in the signal:

$$\widetilde{z}_t = \phi_1(t)\widetilde{z}_{t-1} + \phi_2(t)\widetilde{z}_{t-2} + \sigma(t)\psi_t,$$

where $\phi_1(t) = \phi_1(t+T)$, $\phi_2(t) = \phi_2(t+T)$, and $\sigma(t) = \sigma(t+T)$ are the model parameters. Figure 4 plots the fitted curve (in blue) superimposed on $\{\tilde{z}_t, t = 1, ..., N\}$ (in red). From Figure 4, we note that the PAR(2) model fits the signal reasonably well, confirming the efficiency of our algorithm to estimate the parameters and the validity of Proposition 2 to transform an EPACS signal to cyclostationarity.

E. Hypothesis Test

To illustrate the procedure discussed in Subsection III-C, we tested for constant amplitude given the choices of $m_0(t)$, $R_0(s,t)$, l(t), and g(t), as in Subsection IV-A. To simulate, we considered sample size n = 200 and the parameters $\alpha_1 = 20$,



Fig. 4. Plot of the fitted PAR(2) model (in blue) superimposed on the signal $\{\tilde{z}_t, t = 1, \dots, N\}$ (in red).



Fig. 5. Plot of the variable star R Hydrae during the period 1900–1950. The x-axis plots the days, while the y-axis plots the magnitudes of the light emitted from the star R Hydrae.

 $\alpha_2 = 20/n$, and $\beta = 0, n^{-1}/100, n^{-1}/10, n^{-1}, 10n^{-1}$. The exceedance threshold δ was chosen such that

$$\frac{1}{\lfloor l/2 \rfloor} \|\boldsymbol{p} - \boldsymbol{q}\|_2 = \frac{1}{\lfloor l/2 \rfloor} \sum_{h=1}^{\lfloor l/2 \rfloor} (p_h - q_h)^2 > 10^{-8},$$

which implies $\delta = 10^{-8} \lfloor l/2 \rfloor$. Finally, to calculate the sample squared coherence, $|\mathcal{C}(p, q, M)|^2$, we considered the smoothing window with M = 8. The percentages of rejecting the null hypothesis of $\beta = 0$ against $\beta \neq 0$ and the corresponding 95% binomial confidence intervals were calculated over 500 simulations and are presented in Table II, which suggests that the empirical level of the considered test can be made close to the 5% nominal level.

Noting that the synthetic results are in line with the proposed theories, in the next section we fit an EPACS model to a physical signal from astronomy.

V. ASTRONOMY SIGNAL EXAMPLE: R HYDRAE

A. Data Description

We considered the magnitudes of the light emitted from the variable star R Hydrae during 1900–1950 with an irregular time design consisting of 1086 time points. It is a long period variable star, and its amplitude and period are evolving over time; see [5] and [6]. Figure 5 plots the signal under consideration.

 TABLE II

 Rejection Rates (%) of the Null Hypothesis and the Corresponding 95% Binomial Confidence Intervals for the Parameter

 $\beta = 0, n^{-1}/100, n^{-1}, 10n^{-1}$

$R_0(s,t)$	$\beta = 0$	$\beta = 5 \cdot 10^{-5}$	$\beta = 5 \cdot 10^{-4}$	$\beta = 5 \cdot 10^{-3}$	$\beta = 5 \cdot 10^{-2}$
$\exp(- s-t)$	5.6(1.1, 10.1)	6.9(2.8, 11.0)	34.8 (25.5, 44.1)	96.8 (93.4, 100.0)	100.0
$(1+ s-t)\exp(- s-t)$	$2.9\ (1.0, 4.9)$	$15.8\ (8.7, 22.9)$	$37.6\ (28.1, 47.1)$	$97.9\ (93.6, 100.0)$	100.0
$\min\{s,t\}$	$3.6\ (0.1, 7.1)$	$4.6\ (0.9, 8.3)$	$36.8\ (27.3, 46.3)$	$96.6\ (93.0, 100.0)$	100.0



Fig. 6. Periodogram plot of the deformed physical signal $\{z_t, t = 1, \dots, N\}$.



Fig. 7. Coherent statistic plot of the deformed physical signal $\{z_t, t = 1, \dots, N\}$.

B. Estimation

To proceed with the analysis, we expressed the signal in the form of a trajectory as $\{(t, x_t), t \in \mathbb{D}\}$, where $\mathbb{D} = \{t_1, \ldots, t_n\}$ denotes the set of time points at which the signal was observed and n(= 1086) represents the sample size. Linear models were assumed for both the amplitude and the (instantaneous) period functions:

$$l(t|\beta) = 1 + \beta t \text{ and } g(t|\alpha) = \alpha_2^{-1} \log \left(1 + \alpha_1^{-1} \alpha_2 t\right), \alpha_1 > 0.$$

Using the developed methodology in Subsection III-B, the estimates found were $\hat{\beta} = -1.12 \times 10^{-6}$, $\hat{\alpha}_1 = 419.1$, and $\hat{\alpha}_2 = -1.42 \times 10^{-3}$.

C. Transforming the Signal to Cyclostationarity

Given the parameter estimates, we deformed the signal to have $\{(r_t = g(t|\hat{\alpha}_1, \hat{\alpha}_1), y_t = x_t/l(t|\hat{\beta})), t \in \mathbb{D}\}$ to make it cyclostationary. Because the data points of the deformed signal are not regularly spaced, similar to Subsection IV-C, we interpolated them at regularly spaced (distances of 0.05 units each) time points to have $\{z_t, t = 1, \ldots, N\}$ with N = 875.

Now, to find the period of the deformed signal, we plotted the corresponding periodogram as in Figure 6. The plot displayed a significantly high peak at frequency ≈ 0.052 , suggesting a dominating period of $T \approx 19$. Finally, to confirm that the signal is also cyclostationary, we considered the plot of coherent statistic, $|\mathcal{C}(u-v,0,N)|^2$ [47]; see Figure 7. We observed the most prominent peak at $h \approx 54.6$, which implied that the process was indeed cyclostationary, with a period of approximately 19.



Fig. 8. Plot of the fitted PAR(1) model (in blue) superimposed on the signal $\{\widetilde{z}_t,t=1,\ldots,N\}$ (in red).

D. Model Fitting

Finally, to conclude, we fitted a PAR(1) model to the deformed (regularly spaced) signal after removing its mean $m_z(t) = \frac{1}{L} \sum_{j=0}^{L-1} z_{t+jT}$, t = 1, ..., T, with L denoting the number of cycles present in the signal. Figure 8 plots the signal $\{\tilde{z}_t = z_t - m_z(t), t = 1, ..., N\}$ (in red) along with the fitted curve (in blue); the plot clearly suggests that the PAR(1) model fits the signal $\{\tilde{z}_t, t = 1, ..., N\}$ significantly well. This implies that, to begin with, an EPACS process is "appropriate" to explain the light emitted from the star R Hydrae.

VI. DISCUSSION

In this study, we extended the notion of cyclostationarity to processes that have first- and second-order moments departed from strict periodicities and constant amplitudes and called them EPACS processes. Upon introducing the EPACS processes and discussing a few immediate properties, we provided necessary inference methodologies followed by a synthetic and an astronomical signal example, which exhibited satisfactory numerical results.

We recall from the Introduction that the flexibility of the EPACS processes allows us to model a much larger class of stochastic processes. Therefore, we expect that the developed methodologies must make a significant contribution in modeling a much larger segment of real-life scenarios we encounter. In addition, this article links the EPACS processes back to the classical cyclostationary processes (see Proposition 2), making the fitting of an EPACS process easy and dependent on the existing methods for cyclostationary processes.

In this article, we focused mainly on building the foundation of the EPACS processes and provided semi-parametric estimation and testing procedures. However, other methods to directly estimate the functions m_0 and R_0 could be explored as well, such as kernels and wavelets. Although not studied here, the frequency domain analysis of such processes, EPACS processes with multiple periods, and a multivariate counterpart of the EPACS processes as an extension of multivariate cyclostationary processes can also be explored in the future.

APPENDIX A PROOF OF PROPOSITION 1

We know that a necessary and sufficient condition for a bivariate function $h(s,t): \mathbb{R}^2 \to \mathbb{C}$ to be a valid covariance function of some random signal is that $\{h(s,t), s, t \in \mathbb{R}\}$ is non-negative definite; that is, for every n, for any sequence of real numbers $\{x_1, \ldots, x_n\}$, and for any sequence of complex constants $\{c_1, \ldots, c_n\}$,

$$\sum_{j=1}^{n} \sum_{k=1}^{n} c_j \overline{c_k} h(x_j, x_k) \ge 0$$

holds. Now, because l(t) is strictly positive and g(t) is a strictly increasing function such that g'(t) > 0 for all $t \in \mathbb{R}$, we note that

$$\sum_{j=1}^{n} \sum_{k=1}^{n} c_j \overline{c_k} R_0(x_j, x_k)$$

=
$$\sum_{j=1}^{n} \sum_{k=1}^{n} \frac{c_j}{l(g^{-1}(x_j))} \frac{\overline{c_k}}{l(g^{-1}(x_k))} R\left(g^{-1}(x_j), g^{-1}(x_k)\right) \ge 0.$$

APPENDIX B PROOFS OF THE PROPERTIES OF SECTION II-C

For the proofs below, we denote an EPACS process satisfying (1) by EPACS_X -{ $l(t), g(t), m_0(t), R_0(s, t)$ }.

Property 2.3.1: We note that $E(X_t) = l(t)m_Y(g(t))$ and $cov(X_s, X_t) = l(s)l(t)R_Y(g(s), g(t))$. Therefore, in this scenario, $\{X_t, t \in \mathbb{R}\}$ is $EPACS_X$ - $\{l(t), g(t), m_Y(t), R_Y(s, t)\}$ with $m_Y(t) = m_Y(t+1)$ and $R_Y(s, t) = R_Y(s+1, t+1)$.

Property 2.3.2: Consider two uncorrelated EPACS processes EPACS_X - $\{l(t), g(t), m_{01}(t), R_{01}(s, t)\}$ and EPACS_Y - $\{l(t), g(t), m_{02}(t), R_{02}(s, t)\}$.

This implies that the mean and covariance functions of $\{Z_t = X_t + Y_t, t \in \mathbb{R}\}$ are

 $\mathbf{E}(Z_t) = l(t) \left(m_{01}(g(t)) + m_{02}(g(t)) \right)$

and

$$\operatorname{cov}(Z_s, Z_t) = l(s)l(t) \left(R_{01}(g(s), g(t)) + R_{02}(g(s), g(t)) \right),$$

respectively. That is, $\{Z_t = X_t + Y_t, t \in \mathbb{R}\}$ is EPACS_Z- $\{l(t), g(t), m_{01}(t) + m_{02}(t), R_{01}(s, t) + R_{02}(s, t)\}$.

Property 2.3.3: Let $\{X_t, t \in \mathbb{R}\}$ be EPACS_X - $\{l(t), g(t), m_0(t), R_0(s, t)\}$. Then, the process $\{Y_t = f_t X_t, t \in \mathbb{R}\}$ has mean and covariance functions,

$$\mathbf{E}(Y_t) = e(t)h_{g(t)}l(t)m_0\left(g(t)\right)$$

and

$$\operatorname{cov}(Y_s, Y_t) = e(s)e(t)h_{g(s)}h_{g(t)}l(s)l(t)R_0(g(s), g(t)).$$

This makes $\{Y_t, t \in \mathbb{R}\}$ an EPACS_Y- $\{e(t)l(t), g(t), h_t m_0(t), h_s h_t R_0(s, t)\}$.

Property 2.3.4: We note that

$$\mathsf{E}(Y_t) = \sum_{j=1}^{N} e(t) h_{g(t)}^j l(t) m_0^j(g(t))$$

and

$$\operatorname{cov}(Y_s, Y_t) = \sum_{j=1}^{N} \sum_{k=1}^{N} e(s) h_{g(s)}^j e(t) h_{g(t)}^k l(s) l(t) R_0^{jk}(g(s), g(t)).$$

 $\begin{array}{ll} \text{This implies that } \{Y_t,t\in\mathbb{R}\} \text{ is } \text{EPACS}_Y \hbox{-} \{e(t)l(t),g(t),\\ \sum_{j=1}^N h_t^j m_0^j(t), \sum_{j=1}^N \sum_{k=1}^N h_s^j h_t^k R_0^{jk}(s,t)\}. \\ \text{Property 2.3.5: Let } \{X_t,t\in\mathbb{R}\} \text{ be } \text{EPACS}_X \hbox{-} \{l(t),g(t),\\ \end{array}$

Property 2.3.5: Let $\{X_t, t \in \mathbb{R}\}$ be EPACS_X- $\{l(t), g(t), m_0(t), R_0(s, t)\}$. Then, $\{Y_t, t \in \mathbb{R}\}$ has mean and covariance functions,

$$\mathbf{E}(Y_t) = a(t)l(f(t))m_0(g(f(t)))$$

and

$$cov(Y_s, Y_t) = a(s)a(t)l(f(s))l(f(t))R_0(g(f(s)), g(f(t))).$$

That is, $\{Y_t, t \in \mathbb{R}\}$ is EPACS_X- $\{a(t)(l(f(t)), g(f(t)), m_0(t), R_0(s, t)\}$. The second part of the proposition is a trivial consequence.

APPENDIX C PROOF OF PROPOSITION 3

Let $X_t = l(t)V^{g(t)}P_{g(t)}$ hold for some unitary operator V and some unit-periodic signal $\{P_t, t \in \mathbb{R}\}$. To show that $\{X_t, t \in \mathbb{R}\}$ is EPACS, write $X_t = l(t)Y_{g(t)}$, where $Y_t = V^tP_t$. By writing it this way, we note that $\{Y_t, t \in \mathbb{R}\}$ is cyclostationary (see the 'only if' part of Proposition 7.2 of [9]) and, therefore, $\{X_t, t \in \mathbb{R}\}$ is EPACS by construction.

Now, let $\{X_t, t \in \mathbb{R}\}$ be EPACS. Then, using Proposition 2, the time-deformed process $\{Y_t : Y_{g(t)} = X_t/l(t), t \in \mathbb{R}\}$ is cyclostationary with period one. That is, we have $X_t = l(t)Y_{g(t)}$ with $\{Y_t = \frac{1}{l\{g^{-1}(t)\}}X_{g^{-1}(t)}, t \in \mathbb{R}\}$ cyclostationary. Therefore, there exists a unitary operator V and a unit-periodic signal P_t such that, for every t, $Y_t = V^t P_t$ holds (see the 'if' part of Proposition 7.2 of [9]). Hence, the result follows.

APPENDIX D PROOF OF PROPOSITION 4

We assume that the mean and covariance functions of the signal $\{\mathbf{Z}_t\}$ are such that $\mathrm{E}(Z_t^j) = \mu^j$ and $\mathrm{cov}(Z_s^j, Z_t^k) = \Gamma^{jk}(s-t)$, respectively. Now, if (3) holds, then

$$m(t) = l(t) \sum_{j=-\infty}^{\infty} \mu^j \exp\left(i2\pi jg(t)\right)$$

and

R(s,t)

$$= l(s)l(t)\sum_{j,k=-\infty}^{\infty}\Gamma^{jk}(g(s) - g(t))\exp\left(i2\pi\left(jg(s) + kg(t)\right)\right).$$

Therefore, clearly, $\{X_t, t \in \mathbb{R}\}$ is EPACS with

$$m_0(t) = \sum_{j=-\infty}^{\infty} \mu^j \exp(i2\pi jt)$$

and

$$R_0(s,t) = \sum_{j,k=-\infty}^{\infty} \Gamma^{jk}(s-t) \exp\left(i2\pi(js+kt)\right).$$

On the other hand, if $\{X_t, t \in \mathbb{R}\}$ is EPACS following (1), we express it as (2) and use Fourier representation of $P_t = \sum_{j=-\infty}^{\infty} \widetilde{P_j} \exp(i2\pi jt)$ to get (3), where $Z_t^j = V^t \widetilde{P_j}$. As also mentioned in [9], Z_t^j are orbits of different starting vectors $\widetilde{P_j}$ under the same unitary operator V, and thus they are jointly stationary.

Finally, we remark that the result of Proposition 4 can also be derived from the harmonic series representation, which is valid for cyclostationary processes, given in [54].

References

- L. Eyer and M. G. Genton, "Characterization of variable stars by robust wave variograms: An application to Hipparcos mission," *Astron. Astrophys. Suppl. Ser.*, vol. 136, no. 2, pp. 421–428, 1999.
- [2] C. Koen, "Statistics of OC diagrams and period changes," in *Proc. Light-Time Effect Astrophys.: Causes Cures O-C Diagram*, vol. 335, 2005, p. 25.
- [3] F. Rodler and E. Guggenberger, "Spurious period shifts and changes among variable stars," in *Proc. Light-Time Effect Astrophys.: Causes Cures O-C Diagram*, vol. 335, 2005, p. 115.
- [4] J. D. Hart, C. Koen, and F. Lombard, "An analysis of pulsation periods of long period variable stars," *J. Roy. Stat. Society: Ser. C. (Applied Statistics)*, vol. 56, no. 5, pp. 587–606, 2007.
- [5] A. A. Zijlstra, T. Bedding, and J. Mattei, "The evolution of the Mira variable R Hydrae," *Monthly Notices Roy. Astronomical Soc.*, vol. 334, no. 3, pp. 498–510, 2002.
- [6] M. G. Genton and P. Hall, "Statistical inference for evolving periodic functions," J. Roy. Stat. Soc.: Ser. B (Stat. Methodol.), vol. 69, no. 4, pp. 643–657, 2007.
- [7] W. A. Gardner, "An introduction to cyclostationary signals," in *Cyclostationarity in Communications and Signal Processing*. New York, NY, USA: IEEE Press, 1994, pp. 1–90.
- [8] G. B. Giannakis, "Cyclostationary signal analysis," in *Digital Signal Processing Handbook*, pp. 17–1, 1998.

- [9] H. L. Hurd and A. Miamee, *Periodically Correlated Random Sequences: Spectral Theory and Practice*. Hoboken, NJ, USA: Wiley, 2007, vol. 355.
- [10] J. Antoni, N. Ducleaux, G. NGhiem, and S. Wang, "Separation of combustion noise in IC engines under cyclo-non-stationary regime," *Mech. Syst. Signal Process.*, vol. 38, no. 1, pp. 223–236, 2013.
- [11] D. Abboud, S. Baudin, J. Antoni, D. Rémond, M. Eltabach, and O. Sauvage, "The spectral analysis of cyclo-non-stationary signals," *Mech. Syst. Signal Process.*, vol. 75, pp. 280–300, 2016.
- [12] D. Abboud, J. Antoni, M. Eltabach, and S. Sieg-Zieba, "Angle time cyclostationarity for the analysis of rolling element bearing vibrations," *Measurement*, vol. 75, pp. 29–39, 2015.
- [13] R.-B. Sun, Z.-B. Yang, K. Gryllias, and X.-F. Chen, "Cyclostationary modeling for local fault diagnosis of planetary gear vibration signals," *J. Sound Vib.*, vol. 471, 2020, Art. no. 115175.
- [14] S. Lupenko, N. Lutsyk, and Y. Lapusta, "Cyclic linear random process as a mathematical model of cyclic signals," *Acta Mechanica Et Automatica*, vol. 9, no. 4, pp. 219–224, 2015.
- [15] H. Miao, F. Zhang, and R. Tao, "New statistics of the second-order chirp cyclostationary signals: Definitions, properties and applications," *IEEE Trans. Signal Process.*, vol. 67, no. 21, pp. 5543–5557, Nov. 2019.
- [16] S. C. Olhede and H. Ombao, "Modeling and estimation of covariance of replicated modulated cyclical time series," *IEEE Trans. Signal Process.*, vol. 61, no. 8, pp. 1944–1957, Apr. 2013.
- [17] P. Borgnat, P. Flandrin, and P.-O. Amblard, "Stochastic discrete scale invariance," *IEEE Signal Process. Lett.*, vol. 9, no. 6, pp. 181–184, Jun. 2002.
- [18] H. Saleur and D. Sornette, "Complex exponents and log-periodic corrections in frustrated systems," *J. De Physique I*, vol. 6, no. 3, pp. 327–355, 1996.
- [19] W. A. Gardner, "Signal interception: A unifying theoretical framework for feature detection," *IEEE Trans. Commun.*, vol. 36, no. 8, pp. 897–906, Aug. 1988.
- [20] W. A. Gardner and E. A. Robinson, "Statistical spectral analysis—A nonprobabilistic theory," J. Dyn. Syst., Meas., Control, vol. 111, no. 4, pp. 673–673, 12 1989.
- [21] W. A. Gardner and C. M. Spooner, "Signal interception: Performance advantages of cyclic-feature detectors," *IEEE Trans. Commun.*, vol. 40, no. 1, pp. 149–159, Jan. 1992.
- [22] L. Izzo and A. Napolitano, "Multirate processing of time series exhibiting higher order cyclostationarity," *IEEE Trans. Signal Process.*, vol. 46, no. 2, pp. 429–439, Feb. 1998.
- [23] W. A. Gardner and C. M. Spooner, "The cumulant theory of cyclostationary time-series. I. Foundation," *IEEE Trans. Signal Process.*, vol. 42, no. 12, pp. 3387–3408, Dec. 1994.
- [24] C. M. Spooner and W. A. Gardner, "The cumulant theory of cyclostationary time-series. II. Development and applications," *IEEE Trans. Signal Process.*, vol. 42, no. 12, pp. 3409–3429, Dec. 1994.
- [25] L. Izzo and A. Napolitano, "Higher-order cyclostationarity properties of sampled time-series," *Signal Process.*, vol. 54, no. 3, pp. 303–307, 1996.
- [26] Y.-C. Chen, M.-Y. Cheng, and H.-T. Wu, "Non-parametric and adaptive modelling of dynamic periodicity and trend with heteroscedastic and dependent errors," *J. Roy. Stat. Soc.: Ser. B: Stat. Methodol.*, pp. 651–682, 2014.
- [27] M. S. Taqqu, "A representation for self-similar processes," Stochastic Processes Their Appl., vol. 7, no. 1, pp. 55–64, 1978.
- [28] B. Tsybakov and N. D. Georganas, "Self-similar processes in communications networks," *IEEE Trans. Inf. Theory*, vol. 44, no. 5, pp. 1713–1725, Sep. 1998.
- [29] K.-I. Sato, "Self-similar processes with independent increments," *Probability Theory Related Fields*, vol. 89, no. 3, pp. 285–300, 1991.
- [30] P. Embrechts and M. Maejima, "An introduction to the theory of self-similar stochastic processes," *Int. J. Modern Phys. B*, vol. 14, pp. 1399–1420, 2000.
- [31] M. G. Genton, O. Perrin, and M. S. Taqqu, "Self-similarity and Lamperti transformation for random fields," *Stochastic Models*, vol. 23, no. 3, pp. 397–411, 2007.
- [32] J. C. Pardo, "A brief introduction to self-similar processes," *Department of Mathematical Sciences, University of Bath, United Kingdom*, vol. 9, pp. 305–316, 2007.
- [33] M. Lee, M. G. Genton, and M. Jun, "Testing self-similarity through Lamperti transformations," J. Agricultural, Biol., Environ. Statist., vol. 21, no. 3, pp. 426–447, 2016.
- [34] V. Pipiras and M. S. Taqqu, Stable Non-Gaussian Self-Similar Processes With Stationary Increments. Berlin, Germany: Springer, 2017.

- [35] W. A. Gardner, "Statistically inferred time warping: Extending the cyclostationarity paradigm from regular to irregular statistical cyclicity in scientific data," *EURASIP J. Adv. Signal Process.*, vol. 2018, no. 1, 2018, Art. no. 59.
- [36] A. Napolitano, "Time-warped almost-cyclostationary signals: Characterization and statistical function measurements," *IEEE Trans. Signal Process.*, vol. 65, no. 20, pp. 5526–5541, Oct. 2017.
- [37] K. Mullen, D. Ardia, D. L. Gil, D. Windover, and J. Cline, "DEoptim: An R package for global optimization by differential evolution," *J. Stat. Softw.*, vol. 40, no. 6, pp. 1–26, 2011.
- [38] R Core Team, R: A. Language and Environment for Statistical Computing, R. Foundation for Statistical Computing, Vienna, Austria, 2019, ISBN 3-900051-07-0. [Online]. Available: https://www.R-project.org/
- [39] H. L. Hurd, "An investigation of periodically correlated stochastic processes.," Ph.D. dissertation, Dept. Elect. Eng., Duke University, 1970.
- [40] G. Carter, "Receiver operating characteristics for a linearly thresholded coherence estimation detector," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-25, no. 1, pp. 90–92, Feb. 1977.
- [41] W. Gardner, "Measurement of spectral correlation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-34, no. 5, pp. 1111–1123, Oct. 1986.
- [42] A. Napolitano, "Discrete-time estimation of second-order statistics of generalized almost-cyclostationary processes," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1670–1688, May 2009.
- [43] A. Napolitano and W. A. Gardner, "Algorithms for analysis of signals with time-warped cyclostationarity," in *Proc. 50th Asilomar Conf. Signals, Syst. Comput.*, 2016, pp. 539–543.
- [44] A. Napolitano, "On characterization and application of oscillatory almostcyclostationary processes," in *Proc. 25th Eur. Signal Process. Conf.*, 2017, pp. 1654–1658.
- [45] J. Lamperti, "Semi-stable stochastic processes," *Trans. Amer. Math. Soc.*, vol. 104, no. 1, pp. 62–78, 1962.
- [46] C. M. Spooner, Performance Evaluation of Detectors for Cyclostationary Signals. Univ. of California., Davis, 1988.
- [47] H. L. Hurd and N. L. Gerr, "Graphical methods for determining the presence of periodic correlation," *J. Time Ser. Anal.*, vol. 12, no. 4, pp. 337–350, 1991.
- [48] W. A. Gardner and C. M. Spooner, "Detection and source location of weak cyclostationary signals: Simplifications of the maximum-likelihood receiver," *IEEE Trans. Commun.*, vol. 41, no. 6, pp. 905–916, Jun. 1993.
- [49] P. Bloomfield, H. L. Hurd, and R. B. Lund, "Periodic correlation in stratospheric ozone data," *J. Time Ser. Anal.*, vol. 15, no. 2, pp. 127–150, 1994.
- [50] G. K. Yeung and W. A. Gardner, "Search-efficient methods of detection of cyclostationary signals," *IEEE Trans. Signal Process.*, vol. 44, no. 5, pp. 1214–1223, May 1996.
- [51] K. Benhenni and M. Rachdi, "Nonparametric estimation of average growth curve with general nonstationary error process," *Commun. Statist. - Theory Methods*, vol. 36, no. 6, pp. 1173–1186, 2007.

- [52] G. E. Forsythe, M. A. Malcolm, and C. B. Moler, *Computer Methods for Mathematical Computations*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1977, vol. 259.
- [53] A. E. Dudek, H. Hurd, and W. Wójtowicz, "PARMA models with applications in R," in *Proc. Cyclostationarity: Theory and Methods - II.* Cham, Switzerland, 2015, pp. 131–153.
- [54] W. Gardner and L. Franks, "Characterization of cyclostationary random signal processes," *IEEE Trans. Inf. Theory*, vol. 21, no. 1, pp. 4–14, Jan. 1975.
- [55] A. Napolitano, Cyclostationary Processes and Time Series: Theory, Applications, and Generalizations. Academic Press, 2019.



Soumya Das received the B.Sc. degree in statistics from the Ramakrishna Mission Residential College, Narendrapur, West Bengal, India, in 2015, and the M.Sc. degree in statistics from the Indian Institute of Technology (IIT), Kanpur, India, in 2017.

He is currently working toward the Ph.D. degree in statistics with the King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia. His research interests include cyclostationary processes, statistical signal processing, semiparametric and non-parametric methods, and spatial,

temporal, and spatio-temporal data analyses.



marine science.

Marc G. Genton received the B.Sc. degree in applied mathematics in 1992 and the Ph.D. degree in statistics in 1996 both from the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland.

He is currently a Distinguished Professor of statistics with the King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia. His research interests include statistical analysis, modeling, prediction, and uncertainty quantification of spatiotemporal data, with applications in environmental and climate science, renewable energies, geophysics, and