Accepted version of the article published in IEEE Signal Processing Letters, vol. 29, 2022. DOI: 10.1109/LSP.2022.3149705.
Available at: http://ieeexplore.ieee.org

On Infinite Past Predictability of Cyclostationary Signals

Jaume Riba, *Senior Member, IEEE*, and Marc Vilà, *Student Member, IEEE*.

***Abstract*—This paper explores the asymptotic spectral decomposition of periodically Toeplitz matrices with finite summable elements. As an alternative to polyphase decomposition and other approaches based on Gladyshev representation, the proposed route exploits the Toeplitz structure of cyclic autocorrelation matrices, thus leveraging on known asymptotic results and providing a more direct link to the cyclic spectrum and spectral coherence. As a concrete application, the problem of cyclic linear prediction is revisited, concluding with a generalized Kolmogorov-Szegö theorem on the predictability of cyclostationary signals. These results are finally tested experimentally in a prediction setting for an asynchronous mixture of two cyclostationary pulse-amplitude modulation signals.**

***Index Terms*—Szegö's theorem, cyclostationarity, cyclic Wiener filtering, periodically Toeplitz matrices, spectral coherence.**

NOTATION

- $\mathbf{I}_P$: $P \times P$ identity matrix.
- $|\mathbf{A}|$: determinant of matrix $\mathbf{A}$.
- $[\mathbf{A}]_{r,c}$: (r,c) entry of $R \times C$ matrix $\mathbf{A}$.
- $\mathbf{d}^i$: canonical column N-vector, $\left[\mathbf{d}^i\right]_r \triangleq \delta_{r-i}$.
- $\mathbf{s}^i$: steering column N-vector, $[\mathbf{s}^i]_r \triangleq e^{j2\pi \frac{ri}{N}}/\sqrt{N}$.
- $\mathbf{F}_N \triangleq \left[\mathbf{s}^0, \mathbf{s}^1, \ldots, \mathbf{s}^{N-1}\right]$: unitary $N \times N$ Fourier matrix.
- $\mathbf{a} \odot \mathbf{b}$: element-wise product, $[\mathbf{a} \odot \mathbf{b}]_r \triangleq [\mathbf{a}]_r [\mathbf{b}]_r$.

I. INTRODUCTION

CYCLOSTATIONARY or *periodically correlated* processes are those exhibiting periodically time-variant statistics and arise in different fields of science and engineering [1], [2]. Under appropriate modeling, improved signal processing techniques to exploit them have been proposed in multiple works [3], [4], both for time and frequency domains. In spectrum sensing for cognitive radio, for instance, robust techniques for detecting the presence of users immersed in stationary noise have been designed [5], [6]. Cyclostationarity has also been exploited in [7] to estimate the signal to noise ratio of communications signals from second order statistics, as well as in [8] for linear prediction of geophysical and climatological processes, which are usually strongly phase locked with the daily/seasonal cycle. Further applications can be found in [9].

The spectral decomposition of signal statistics (especially its rank) plays an important role in the analysis of detection and estimation problems, as studied in [10]–[13]. In this sense, preliminary works such as [5], [7] root on the fact that cyclostationary signals span a reduced-dimension subspace structure compared to the one given by their overall spectral support, with implications also on rate-distortion theory [14].

Furthermore, theoretical asymptotic bounds are useful to evaluate the performance of signal processing methods obtained with a finite amount of data [15]. In the context of stationary signals, *Szegö's theorem* [16], [17], which deals with the asymptotic eigenanalysis of Toeplitz matrices, is commonly used to characterize this limiting performance. It has the additional advantage of unveiling intimate structures and physical insights in the frequency domain, which finds application on information theoretic learning [18], [19].

Motivated by the aforementioned ideas, the objective of this paper is to provide a succinct and novel study on the asymptotic eigenanalysis of *periodically Toeplitz matrices* exhibited by the statistics of cyclostationary signals [20]. Instead of using the complete Gray's machinery [17], the Toeplitz structure of cyclic autocorrelation matrices is explicitly exploited, resulting in a general Theorem 1 for cyclostationary signals. This approach is an alternative to the polyphase decomposition from [14], [21] or other techniques based on Gladyshev representation [22], [23], which transforms a cyclostationary scalar time series into a stationary vector-valued stochastic process [24], [25]. The theoretical results obtained are applied afterwards to the estimation problem of linear prediction, culminating in Theorem 2 for one-step linear prediction of cyclostationary signals. These findings extend Kolmogorov-Szegö formula without the need for its Wiener-Masani complete extension for vector processes [26], providing new insights and a simpler connection to the cyclic spectrum and spectral coherence.

II. PRELIMINARIES AND PURPOSE

Let $x(n) \in \mathbb{C}$ be a discrete-time cyclostationary signal with integer cycle period P and periodic autocorrelation function $R(n,m) \triangleq \mathrm{E}\left[x(n+m)x^*(n)\right] = R(n+P,m)$. The cyclic spectrum is defined as

$$S^k(f) \triangleq \sum_m R^k(m) e^{-j2\pi m f}, \quad k = 0, 1, \ldots, P-1, \quad (1)$$

where $R^k(m) \triangleq \frac{1}{P}\sum_{n=0}^{P-1} R(n,m) e^{-j2\pi \frac{nk}{P}}$ is the cyclic autocorrelation function. Let $\mathbf{x}_N(n) \triangleq [x(n), x(n-1), \ldots, x(n-N+1)]^T$ be a run-time vector of length N. The $N \times N$ Hermitian non-negative definite autocorrelation matrix of $\mathbf{x}_N(n)$ and its time-averaged version are given by

$$\mathbf{R}_N(n) \triangleq \mathrm{E}\left[\mathbf{x}_N(n)\mathbf{x}_N^H(n)\right], \quad \bar{\mathbf{R}}_N \triangleq \frac{1}{P}\sum_{n=0}^{P-1} \mathbf{R}_N(n), \quad (2)$$

This work has been supported by the Spanish Ministry of Science and Innovation through project RODIN (PID2019-105717RB-C22 / MCIN / AEI / 10.13039/501100011033). Authors are within Signal Processing and Communications group (SPCOM) (Signal Theory and Communications Department) at Technical University of Catalonia (UPC).

respectively, with entries $[\mathbf{R}_N(n)]_{r,c} = R(n-c, c-r)$ and $[\bar{\mathbf{R}}_N]_{r,c} = R^0(c-r)$. Throughout this section and the next one, we will assume, without loss of generality, that $N \triangleq LP$, for some integer $L \geq 1$. Matrix $\bar{\mathbf{R}}_{LP}$, which has a Toeplitz structure (as it happens with $\mathbf{R}_{LP}(n)$ for stationary signals), represents the relevant second order statistics in applications in which the cyclostationarity of the signal of interest is not exploited. Note that $\bar{\mathbf{R}}_{LP}$ can be seen as an averaged autocorrelation matrix: $\bar{\mathbf{R}}_{LP} = \mathrm{E}_\epsilon\left[\mathbf{R}_{LP}(n+\epsilon)\right]$, where the unknown delay parameter ϵ is assumed random, having the least informative prior for the set $\{0, 1, \ldots, P-1\}$ (*i.e.* uniform). The sequence of *stationarized* matrices $\bar{\mathbf{R}}_{LP}$ for increasing values of L becomes equivalent to a sequence of circulant matrices, under mild conditions such as sufficiently large L and the continuity of the spectrum $S^0(f)$, which guarantees the absolute summability of column elements of $\bar{\mathbf{R}}_{LP}$ [17]. This implies that $\bar{\mathbf{R}}_{LP}$ asymptotically diagonalizes within the unitary Fourier matrix, *i.e.* $\lim_{L\to\infty} \bar{\mathbf{R}}_{LP} = \mathbf{F}_{LP}\mathbf{D}^0_{LP}\mathbf{F}^H_{LP}$. As a result, its eigenvalues converge to uniform samples of $S^0(f)$ in the range $0 \leq f < 1$: $[\mathbf{D}^0_{LP}]_{r,r} = S^0\left(\frac{r}{LP}\right)$ with $r = 0, 1, \ldots, LP-1$. This fact is typically used to reveal the asymptotic behavior of signal processing techniques when they handle a large amount of data samples.

In applications where cyclostationarity is exploited, more informative second order statistics are captured by $\mathbf{R}_{LP}(n)$, which is *P-periodically Toeplitz* (or *P-Toeplitz*, for short [27]), determined by $[\mathbf{R}_{LP}(n)]_{r,c} = [\mathbf{R}_{LP}(n)]_{r+P,c+P}$. The goal of this paper is thus understanding the asymptotic structure of its spectral decomposition as $L \to \infty$, in order to unveil the ultimate performance of *synchronized* signal processing methods as the size of data past is let to grow. This route is provided in the sequel with a theoretic development of spectral decomposition, followed by an application to the relevant case of linear prediction.

III. Asymptotic spectral decomposition of periodically Toeplitz matrices

Let us write the eigenequation of a $LP \times LP$, P-Toeplitz matrix as follows [28]:

$$\mathbf{R}_{LP}(n)\mathbf{q}_{l,p}(n) = \lambda_{l,p}(n)\mathbf{q}_{l,p}(n), \qquad \begin{matrix} 0 \leq l \leq L-1 \\ 0 \leq p \leq P-1 \end{matrix} . \quad (3)$$

The LP eigenpairs are identified using the tuple of subscripts (l,p), with the ultimate purpose of studying the case of $L \to \infty$. The matrix-vector product in (3) is explicitly expressed as:

$$\sum_{c=0}^{LP-1} R\left(n-c, c-r\right) q_{l,p}(n,c) = \lambda_{l,p}(n) q_{l,p}(n,r). \quad (4)$$

Since $R(n,m)$ is P-periodic in n, it admits the Fourier series expansion $R(n,m) = \sum_{k=0}^{P-1} R^k(m) e^{j2\pi\frac{kn}{P}}$, so (4) becomes:

$$\sum_{k=0}^{P-1} e^{j2\pi\frac{kn}{P}} \sum_{c=0}^{LP-1} R^k(c-r) q_{l,p}(n,c) e^{-j2\pi\frac{kc}{P}} = \lambda_{l,p}(n) q_{l,p}(n,r). \quad (5)$$

We convert (5) back into matrix form:

$$\sum_{k=0}^{P-1} e^{j2\pi\frac{kn}{P}} \mathbf{R}^k_{LP}\left(\mathbf{q}_{l,p}(n) \odot \sqrt{LP}\mathbf{s}^{-kL}\right) = \lambda_{l,p}(n)\mathbf{q}_{l,p}(n), \quad (6)$$

where, clearly, $\left[\mathbf{R}^k_{LP}\right]_{r,c} \triangleq R^k(c-r)$ are Toeplitz matrices. Therefore, for large L, all of them asymptotically diagonalize under the unitary Fourier transform:

$$\mathbf{R}^k_{LP} \overset{L\to\infty}{=} \mathbf{F}_{LP}\mathbf{D}^k_{LP}\mathbf{F}^H_{LP}, \quad \left[\mathbf{D}^k_{LP}\right]_{r,r} = S^k\left(\tfrac{r}{LP}\right), \quad (7)$$

for which (6) becomes

$$\sqrt{LP}\sum_{k=0}^{P-1} e^{j2\pi\frac{kn}{P}} \mathbf{D}^k_{LP}\mathbf{F}^H_{LP}\left(\mathbf{q}_{l,p}(n) \odot \mathbf{s}^{-kL}\right) = \lambda_{l,p}(n)\mathbf{F}^H_{LP}\mathbf{q}_{l,p}(n). \quad (8)$$

Up until this point, we have used known asymptotic limits to yield (8). We now state the following main theorem:

Theorem 1. *(On large periodically Toeplitz matrices) The asymptotic eigenvectors of a P-Toeplitz matrix as defined in* (3) *are given by the P-periodic (in n) orthonormal vectors*

$$\mathbf{q}_{l,p}(n) \overset{L\to\infty}{=} \sum_{q=0}^{P-1} b_{l,p}(q) e^{-j2\pi\frac{qn}{P}} \mathbf{s}^{l+qL}, \quad (9)$$

for some particular unit-energy complex sequences $\{b_{l,p}(q)\}_{q=0,1,\ldots,P-1}$, and their associated eigenvalues are asymptotically time-invariant, i.e. $\lambda_{l,p}(n) \overset{L\to\infty}{=} \lambda_{l,p}$.

Proof: The proof is based on checking that (8) is fulfilled by the structure given in (9). Replacing q by r in (9) and plugging it into the right hand side of (8) we have

$$\lambda_{l,p}(n)\mathbf{F}^H_{LP}\left(\sum_{r=0}^{P-1} b_{l,p}(r) e^{-j2\pi\frac{rn}{P}} \mathbf{s}^{l+rL}\right) = \lambda_{l,p}(n)\sum_{r=0}^{P-1} b_{l,p}(r) e^{-j2\pi\frac{rn}{P}} \mathbf{d}^{l+rL}, \quad (10)$$

where we have used that $\mathbf{F}^H_N\mathbf{s}^i = \mathbf{d}^i$. Now, replacing q by c in (9) and plugging it in the left hand side of (8) yields

$$\sum_{k=0}^{P-1} e^{j2\pi\frac{kn}{P}} \mathbf{D}^k_{LP}\mathbf{F}^H_{LP}\left(\sum_{c=0}^{P-1} b_{l,p}(c) e^{-j2\pi\frac{cn}{P}} \mathbf{s}^{l+(c-k)L}\right) = \sum_{c=0}^{P-1} b_{l,p}(c)\sum_{k=0}^{P-1} e^{j2\pi\frac{(k-c)n}{P}} \mathbf{D}^k_{LP}\mathbf{d}^{l+(c-k)L}. \quad (11)$$

To force equality between (10) and (11) for a given l, since the ones in $\mathbf{d}^{l+rL}$ are spaced by L samples for different values of r, it suffices to set $c-k \equiv r$ in (11). By doing so and using (7), we obtain the following P-dimensional eigenequation:

$$\sum_{c=0}^{P-1} b_{l,p}(c) e^{-j2\pi\frac{rn}{P}} S^{c-r}\left(\tfrac{l+rL}{LP}\right) = \lambda_{l,p}(n) b_{l,p}(r) e^{-j2\pi\frac{rn}{P}}$$

$$\sum_{c=0}^{P-1} b_{l,p}(c) S^{c-r}\left(\tfrac{l}{N} + \tfrac{r}{P}\right) = \lambda_{l,p}(n) b_{l,p}(r) \quad (12)$$

$$\mathbf{S}_P\left(\tfrac{l}{N}\right)\mathbf{b}_{l,p} = \lambda_{l,p}\mathbf{b}_{l,p}, \quad (13)$$

where $\mathbf{S}_P(f) \in \mathbb{C}^{P\times P}$ is the Hermitian non-negative definite spectral correlation matrix given by

$$\left[\mathbf{S}_P\left(f\right)\right]_{r,c} = S^{c-r}\left(f + \tfrac{r}{P}\right), \quad (14)$$

and $\mathbf{b}_{l,p} \in \mathbb{C}^P$ are the eigenvectors of the spectral correlation that define the P-length sequences in (9): $\left[\mathbf{b}_{l,p}\right]_q = b_{l,p}(q)$.

Notice that, since the common term $e^{-j2\pi rn/P}$ has been removed on both sides of (12), the dependence of $\lambda_{l,p}$ on n

has disappeared in (13), since both $\mathbf{S}_P(f)$ and $\mathbf{b}_{l,p}$ are time-invariant. Thus, while the eigenvectors are time-variant from (9), the eigenvalues are instead time-invariant. ■

Thm. 1 generalizes the case explored in [7] where the noise subspace of the rank-one spectral correlation matrices of oversampled pulse-amplitude modulation (PAM) signals was exploited as a means to obtain unbiased estimates of the in-band noise power using only second order statistics.

IV. Predictability of cyclostationary signals

As a specific application of the main result given in Thm. 1, let us consider the problem of one-step prediction based on N past samples of the signal. In particular, we focus on the minimum mean square error (MMSE) given by

$$\xi_N(n) \triangleq \min_{\mathbf{h}(n)} \mathrm{E}\left[\left|x(n) - \mathbf{h}^H(n)\mathbf{x}_N(n-1)\right|^2\right]. \quad (15)$$

This is a well-known Wiener filtering problem, referred to as *cyclic Wiener filtering* for cyclostationary signal processing [29], which results in a periodic predictor filter $\mathbf{h}(n)$ that implements a synchronous processing by using the prior knowledge of the cycle period of the signal under analysis.

We are interested in the infinite past predictability of such a signal, *i.e.* $\xi_\infty \triangleq \lim_{N\to\infty}\xi_N(n)$. While classical linear prediction theory for stationary signals provides an elegant answer to this question through a measure of spectral flatness [28], we aim at generalizing the result by providing a predictability measure in terms of the cyclic spectrum, thus giving insights on the achievable predictability gain of synchronous vs. asynchronous signal processing as a by-product.

The well-known normal equations [28] applied to the general cyclostationary case lead to

$$\xi_N(n) = R(n,0) - \mathbf{r}^H(n)\mathbf{R}_N^{-1}(n-1)\mathbf{r}(n), \quad (16)$$

with $\mathbf{r}(n) \triangleq \mathrm{E}\left[\mathbf{x}_{N-1}(n-1)x^*(n)\right]$. Since matrix $\mathbf{R}_{N+1}(n)$ admits the block partitioning

$$\mathbf{R}_{N+1}(n) = \begin{bmatrix} R(n,0) & \mathbf{r}^H(n) \\ \mathbf{r}(n) & \mathbf{R}_N(n-1) \end{bmatrix}, \quad (17)$$

its determinant can be expressed as:

$$|\mathbf{R}_{N+1}(n)| = \left(R(n,0) - \mathbf{r}^H(n)\mathbf{R}_N^{-1}(n-1)\mathbf{r}(n)\right)|\mathbf{R}_N(n-1)|, \quad (18)$$

for which, using (16) and (18), the MMSE becomes:

$$\xi_N(n) = \frac{|\mathbf{R}_{N+1}(n)|}{|\mathbf{R}_N(n-1)|}. \quad (19)$$

Given (19), we can now state the following theorem:

Theorem 2. *(Generalized Kolmogorov-Szegö for cyclostationary signals) A lower bound on the MMSE of one-step prediction of a cyclostationary signal of cycle period P is given by*

$$\xi_\infty^{(P)} = \exp\int_0^{1/P} \ln|\mathbf{S}_P(f)|\,df \le \xi_N(n), \quad (20)$$

with $\mathbf{S}_P(f)$ defined as in (14). *It is asymptotically achieved for infinite past samples ($N\to\infty$), in which case the prediction error becomes stationary. The synchronous vs. asynchronous prediction gain is given by*

$$g \triangleq \frac{\xi_\infty^{(P)}/P_x}{\xi_\infty^{(1)}/P_x} = \exp\int_0^{1/P}\ln|\mathbf{C}_P(f)|\,df \le 1, \quad (21)$$

where P_x is the signal power, and the spectral coherence matrix *is defined as:*

$$\left[\mathbf{C}_P(f)\right]_{r,c} \triangleq \frac{S^{c-r}\left(f+\frac{r}{P}\right)}{S^0\left(f+\frac{r}{P}\right)}. \quad (22)$$

The equality in (21) *is achieved iff the signal is stationary.*

Proof: From Thm. 1, the eigenvalues of $\mathbf{R}_{N+1}(n)$ and $\mathbf{R}_N(n-1)$ converge to the identical time-invariant $\lambda_{l,p}$ in (13), for which the denominator of (19) behaves as follows:

$$\begin{aligned}\lim_{N\to\infty}|\mathbf{R}_N(n-1)| &= \lim_{L\to\infty}\prod_{l=0}^{L-1}\prod_{p=0}^{P-1}\lambda_{l,p} \\ &= \lim_{L\to\infty}\prod_{l=0}^{L-1}\left|\mathbf{S}_P\left(\tfrac{l}{N}\right)\right|. \end{aligned} \quad (23)$$

Then, by taking the power $1/N$ and the logarithm,

$$\begin{aligned}\lim_{N\to\infty}\ln\left(|\mathbf{R}_N(n-1)|^{1/N}\right) &= \lim_{L\to\infty}\tfrac{1}{LP}\sum_{l=0}^{L-1}\ln\left|\mathbf{S}_P\left(\tfrac{l}{LP}\right)\right| \\ = \tfrac{1}{P}\int_0^1 \ln\left|\mathbf{S}_P\left(\tfrac{f}{P}\right)\right|df &= \int_0^{1/P}\ln|\mathbf{S}_P(f)|\,df \triangleq J_P, \end{aligned} \quad (24)$$

which is the result of expressing the asymptotic average in l as an integral. Therefore, $\lim_{N\to\infty}|\mathbf{R}_N(n-1)|^{1/N} = e^{J_P}$.

With similar reasoning, for the numerator of (19) we have

$$\lim_{N\to\infty}|\mathbf{R}_{N+1}(n)|^{1/(N+1)} = e^{J_P}, \quad (25)$$

for which (19) becomes:

$$\xi_\infty^{(P)} = \lim_{N\to\infty}\xi_N(n) = \lim_{N\to\infty}\frac{e^{((N+1)J_P)}}{e^{(NJ_P)}} = e^{J_P}, \quad (26)$$

yielding (20). Regarding gain g in (21), we compute it as:

$$\begin{aligned} g = \frac{\xi_\infty^{(P)}}{\xi_\infty^{(1)}} &= \frac{\exp\int_0^{1/P}\ln|\mathbf{S}_P(f)|\,df}{\exp\int_0^1 \ln S^0(f)df} = \\ &\exp\int_0^{1/P}\left(\ln|\mathbf{S}_P(f)| - \sum_{r=0}^{P-1}\ln S^0\left(f+\tfrac{r}{P}\right)\right)df, \end{aligned} \quad (27)$$

where we have changed the limits of integration of the denominator. Notice that the second term in the integral is simply the sum of the diagonal elements of $\mathbf{S}_P(f)$ or, equivalently, the determinant of $\bar{\mathbf{S}}_P(f) \triangleq \mathbf{S}_P(f)\odot\mathbf{I}_P$. Thus, by splitting this matrix into its square roots and using properties of the logarithm and the determinant, we complete the proof:

$$\begin{aligned} g &= \exp\int_0^{1/P}\left(\ln|\mathbf{S}_P(f)| + \ln\left|\bar{\mathbf{S}}_P^{-1/2}(f)\bar{\mathbf{S}}_P^{-1/2}(f)\right|\right)df \\ &= \exp\int_0^{1/P}\ln\left|\bar{\mathbf{S}}_P^{-1/2}(f)\mathbf{S}_P(f)\bar{\mathbf{S}}_P^{-1/2}(f)\right|df \\ &= \exp\int_0^{1/P}\ln|\mathbf{C}_P(f)|\,df. \end{aligned} \quad (28)$$

■

Kolmogorov-Szegö theorem [28] is thus generalized in the following sense. While for stationary signals having high predictability requires exhibiting (exponentially) deep nulls on the spectrum, for cyclostationary ones it suffices to display spectral correlation matrices with small determinant. In other words, the smaller the determinant of the spectral correlation matrices is without having spectral nulls (high condition number), the higher the gain of synchronous processing becomes. Additionally, the asymptotic stationarity of the prediction error stated in Thm. 2, coming from the time-invariance of the eigenvalues stated in Thm. 1, adds up to its well-known asymptotic whiteness property.

Finally, we may relate Thm. 2 with other noteworthy works in the literature, such as [6]. Within it, the authors prove the *generalized likelihood-ratio test (GLRT)* and the *locally most powerful invariant test (LMPIT)* for detecting cyclostationarity vs. stationarity asymptotically reduce to functionals of coherence matrices expressed in terms of the *Loève spectrum*. As a result, these detectors become asymptotically invariant to linear filtering. It is remarkable that identical invariant second order statistics emerge on the apparently different problem of evaluating the gain in one-step error prediction power by exploiting cyclostationarity studied here.

V. An illustration: mixture of PAM signals

We next test the validity of the previous theoretical results by considering the following asynchronous mixture of two cyclostationary PAM signals of cycle period $P = 2$:

$$x(n) = x_1(n - 2\varepsilon) + x_2(n - 2(\varepsilon - \delta)) + w(n), \qquad (29)$$

where $x_i(n) = \sum_k a_i(k)p(n - 2k)$ for $i \in \{1, 2\}$ with $R_{a_i a_{i'}}(m) = \delta_{i-i'}\delta_m$ and independent stationary additive noise such that $R_{ww}(m) = \sigma^2\delta_m$. A 100% excess bandwidth pulse $p(n) = (\sqrt{8}/\pi)(-1)^n/(1 - 4n^2)$ is considered, with Fourier transform $P(f) = \sqrt{2}\cos(\pi f)$, which is clearly non-null for $-0.5 < f < 0.5$. It can be shown that

$$S^k(f) = e^{-j2\pi k\varepsilon}\cos(2\pi k\delta)\, P(f + k/2)\, P^*(f) + \sigma^2, \qquad (30)$$

for $k \in \{0, 1\}$. This model permits a full transition from stationary $x(n)$ ($\delta = 0.5$) to full cyclostationarity ($\delta = 0$) with rank-one spectral coherence matrices [7]. Therefore, parameter $\delta \in [0, 0.5]$ will be used to check the convergence of predictability to the theoretical lower bound for large data, and parameter $\varepsilon \in [0, 1)$ will be uniformly distributed in the setting of different experiments in order to show the asymptotic stationarity of the prediction error.

Fig. 1 depicts the prediction error power as a function of N for different values of δ, showing its convergence to the corresponding asymptotic lower bounds. For moderate values of N, the prediction error power depends on the values of ε, as seen from the width of the plotted lines, which represents a full sweep across the full range of ε values. On the contrary, this dependency vanishes for large N, which confirms its asymptotic stationarity.

Fig. 2 shows the asymptotic predictability as a function of the noise variance. If δ decreases, the process exhibits a higher degree of cyclostationarity and the prediction error power floor for $\sigma^2 \to 0$ decreases. When $\delta = 0$ the cyclostationarity is maximum, yielding rank-one plus identity spectral correlation matrices and thus autocorrelation matrices with $N/2$ eigenvalues (asymptotically) equal to the noise power. Therefore, the predictability value becomes dominated by σ^2. As a result, the asymptotic prediction error power converges to zero for $\sigma^2 \to 0$, as seen by the lack of error floor in the $\delta = 0$ curve.

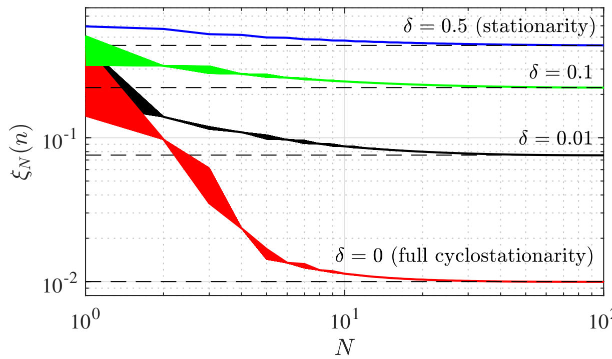

Fig. 1. Prediction error power as a function of N for different values of δ and the full range of ε values. The asymptotic lower bounds are plotted as well with dashed lines. Noise power is set to $\sigma^2 = 5 \times 10^{-5}$.

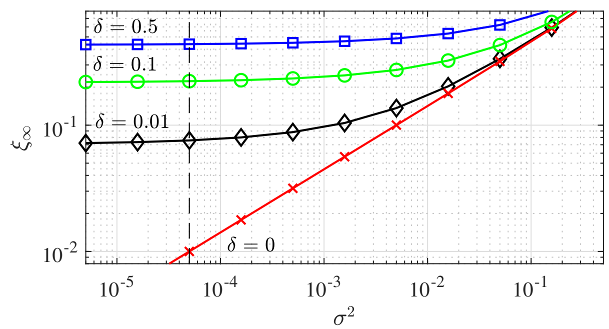

Fig. 2. Asymptotic predictability for different values of δ as a function of the noise power σ^2.

VI. Discussion and future work

Kolmogorov-Szegö theorem for stationary signals, which requires an integral operator of the spectrum logarithm, has been extended to cyclostationary signals, involving an integral operator of the determinant logarithm of spectral correlation matrices. The presented succinct derivation (compared with more general Wiener-Masani extensions and polyphase decompositions) is the result of applying known asymptotic results to Toeplitz cyclic autocorrelation matrices after re-stating the spectral decomposition problem of P-Toeplitz matrices. The obtained lower bounds on the one-step predictability have been validated for different degrees of cyclostationarity.

As highlighted in [27], P-Toeplitz matrices whose size is a multiple of P (such as in Section III) become *block-Toeplitz* with block-size P. This structure has been widely used as a tool in detection problems dealing with cyclostationary signals [6], [11], [30]. Future lines of research can aim at revealing further relationships between the presented work and the aforementioned literature. The ideas developed can also provide new insights onto other signal processing problems, such as data compression [14], [31].

References

[1] E. Serpedin, F. Panduru, I. Sarı, and G. B. Giannakis, "Bibliography on cyclostationarity," *Signal Processing*, vol. 85, no. 12, pp. 2233–2303, Dec 2005.

[2] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Signal Process.*, vol. 86, no. 4, pp. 639–697, Apr. 2006.

[3] A. Napolitano, "Cyclostationarity: New trends and applications," *Signal Processing*, vol. 120, pp. 385–408, Mar 2016.

[4] ——, "Cyclostationarity: Limits and generalizations," *Signal Processing*, vol. 120, pp. 323–347, Mar 2016.

[5] J. Riba, J. Font-Segura, J. Villares, and G. Vázquez, "Frequency domain GLR detection of a second-order cyclostationary signal over fading channels," *IEEE Trans. Signal Process.*, vol. 62, no. 8, pp. 1899–1911, Apr. 2014.

[6] D. Ramírez, P. J. Schreier, J. Vía, I. Santamaría, and L. L. Scharf, "Detection of multivariate cyclostationarity," *IEEE Trans. Signal Process*, vol. 63, no. 20, pp. 5395–5408, Oct. 2015.

[7] J. Riba, J. Villares, and G. Vázquez, "A nondata-aided SNR estimation technique for multilevel modulations exploiting signal cyclostationarity," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5767–5778, Nov. 2011.

[8] K.-Y. Kim, "Statistical prediction of cyclostationary processes," *Journal of Climate*, vol. 13, no. 6, pp. 1098–1115, 15 Mar. 2000.

[9] J. Antoni, "Cyclostationarity by examples," *Mechanical Systems and Signal Processing*, vol. 23, no. 4, pp. 987–1036, May 2009.

[10] O. Besson, S. Kraut, and L. L. Scharf, "Detection of an unknown rank-one component in white noise," *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2835–2839, Jul. 2006.

[11] D. Ramírez, G. Vázquez-Vilar, R. López-Valcarce, J. Vía, and I. Santamaría, "Detection of rank-P signals in cognitive radio networks with uncalibrated multiple antennas," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3764–3774, Aug. 2011.

[12] J. Sala-Álvarez, G. Vázquez-Vilar, R. López-Valcarce, S. Sedighi, and A. Taherpour, "Multiantenna GLR detection of rank-one signals with known power spectral shape under spatially uncorrelated noise," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6269–6283, Dec. 2016.

[13] M. R. Mahmoudi, M. H. Heydari, Z. Avazzadeh, and K.-H. Pho, "Goodness of fit test for almost cyclostationary processes," *Digital Signal Processing*, vol. 96, p. 102597, Jan 2020.

[14] A. Kipnis, A. J. Goldsmith, and Y. C. Eldar, "The distortion rate function of cyclostationary gaussian processes," *IEEE Trans. Inf. Theory*, vol. 64, no. 5, pp. 3810–3824, May. 2018.

[15] M. R. Mahmoudi, M. H. Heydari, and Z. Avazzadeh, "On the asymptotic distribution for the periodograms of almost periodically correlated (cyclostationary) processes," *Digital Signal Processing*, vol. 81, pp. 186–197, Oct 2018.

[16] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*. Berkeley: Univ. Calif. Press, 1958.

[17] R. Gray, *Toeplitz and Circulant Matrices: A review*. Foundations and trends in Communications and Information Theory, 2006.

[18] D. Ramírez, J. Vía, I. Santamaría, and P. Crespo, "Entropy and Kullback-Leibler divergence estimation based on Szegö's theorem," in *European Signal Processing Conference, EUSIPCO*, 2009.

[19] F. de Cabrera and J. Riba, "Squared-loss mutual information via high-dimension coherence matrix estimation," in *IEEE Int. Conf. on Acoust., Speech and Signal Process. (ICASSP)*, 2019, pp. 5142–5146.

[20] M. Chakraborty, "An efficient algorithm for solving general periodic Toeplitz systems," *IEEE Transactions on Signal Processing*, vol. 46, no. 3, pp. 784–787, 1998.

[21] M. Vetterli, J. Kovacevic, and V. K. Goyal, *Foundations of Signal Processing*, U. Cambridge, Ed. Cambridge Univ. Pres, 2014.

[22] E. D. Gladyshev, "Periodically correlated random sequences," *Soviet Math. Dokl.*, vol. 2, pp. 285–388, 1961.

[23] H. L. Hurd and A. Miamee, *Periodically Correlated Random Sequences: Spectral Theory and Practice (Wiley Series in Probability and Statistics)*. USA: Wiley-Interscience, 2007.

[24] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time-domain approach," *IEEE Trans. Inf. Theory*, vol. 40, pp. 340–350, Mar. 1994.

[25] R. López-Valcarce and S. Dasgupta, "Blind channel equalization with colored sources based on second-order statistics: A linear prediction approach," *IEEE Trans. Signal Process.*, vol. 49, no. 9, pp. 2050–2059, Sep. 2001.

[26] N. Wiener and P. R. Masani, "The prediction theory of multivariate stochastic processes," *Acta Mathematica*, vol. 98, pp. 111–150, 1957.

[27] M. J. C. Gover and S. Barnett, "Inversion of Toeplitz matrices which are not strongly non-singular," *IMA Journal of Numerical Analysis*, vol. 5, no. 1, pp. 101–110, 01 1985.

[28] P. P. Vaidyanathan, *The Theory of Linear Prediction*. Morgan and Claypool Publishers, 2008.

[29] W. A. Gardner, "Cyclic Wiener filtering: Theory and method," *IEEE Trans. Commun.*, vol. 41, no. 1, pp. 151–163, Jan. 1993.

[30] D. Ramírez, J. Vía, I. Santamaría, and L. L. Scharf, "Locally most powerful invariant tests for correlation and sphericity of Gaussian vectors," *IEEE Trans. Inf. Theory*, vol. 59, no. 4, pp. 2128–2141, Apr. 2013.

[31] B. A. Francis and S. Dasgupta, "Signal compression by subband coding," *Autom.*, vol. 35, pp. 1895–1908, 1999.