On the Prediction of a Class of Wide-Sense Stationary Random Processes

Juan Miguel Medina and Bruno Cernuschi-Frías, *Senior Member, IEEE*

***Abstract*—We prove that under suitable conditions, a multi-band wide sense stationary stochastic process can be linearly predicted at time t with arbitrarily small error using past samples taken at uniform rate. This result generalizes previous similar results for band-limited signals. Moreover we prove that the prediction problem from uniform past samples is equivalent to a disjoint translates condition on the spectrum together with the divergence of a logarithmic integral. We also show that, for the band-limited case, under similar conditions, non uniform samples can be taken.**

***Index Terms*—Prediction, stationary random processes, statistical signal processing.**

I. Introduction

We consider the classical problem of linearly predicting the current value of a continuous time wide-sense stationary (w.s.s.) random process X_t, over a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, in terms of its past samples. We study conditions to predict with arbitrarily small error the current value at an instant $t+\tau$, $\tau \geq 0$, given past samples taken at uniform rate $X_{t-1}, X_{t-2}, \ldots$.

From the classical Szegö theorem [13], one gets that [11, Ch. 2, p. 80] or [10, Ch. IV]

$$\int_{\mathbb{R}} \log\left(\frac{d\mu_\pi}{d\lambda}(x)\right)(1+x^2)^{-1}dx = -\infty \tag{1}$$

where $d\mu/d\lambda$ is the Radon–Nykodym derivative respect the Lebesgue measure λ of the spectral measure μ of the process X_t, is a necessary and sufficient condition for the prediction of a continuous time wide sense stationary process at an instant t [10]. In other words, this condition ensures that given $s \in \mathbb{R}$, $\overline{\text{span}}\{X_t\}_{t\leq s} = \overline{\text{span}}\{X_t\}_{t\in\mathbb{R}}$ (the closed linear span with respect to the $L^2(\Omega, \mathcal{F}, \mathbb{P})$ norm). To find the predictor may become rather involved in the general case. Some solutions for particular cases are known. In particular, one may be interested in making the prediction using *only* past samples taken at uniform rate, and not using the whole past. In Hilbert space language this is equivalent to find conditions for $\overline{\text{span}}\{X_k\}_{t\in\mathbb{R}} = \overline{\text{span}}\{X_k\}_{k\in\mathbb{Z}_{<0}}$ to hold. For band-limited processes with absolutely continuous spectral measure i.e., processes with spectral density, Wainstein and Zubakov [22] proved that if the sampling rate is increased at least three times above the Nyquist rate. A band-limited process can be predicted with arbitrarily small error from its past samples using a "universal" formula for the predictor. A better result in this direction is [7], where a similar predictor is constructed when the samples are taken at twice the Nyquist rate. This sequence of predictors converges with exponential rate. However, despite it could be more difficult to find explicit coefficients for the predictor, more general conditions are, for example, given in [3]. Particular cases of this result are those of [6] and [17], where it is shown that a sufficient condition is to take the samples at an arbitrarily rate greater than Nyquist's rate. In [15], sufficient conditions for the predictability of a multi-band process in terms of uniform past samples are given, as well as a simple expression for the prediction error. The result of [15] contains some of the previous results as particular cases. Again, this result assumes the existence of the spectral density of the process and the proofs are in the same spirit of the work of Beaty and Dodson [1] which extends the classical Shannon–Nyquist theorem for deterministic signals to the non band-limited case.

We present a result that generalizes [15], giving an alternative proof. In this work we drop the hypothesis of the absolute continuity of the spectral measure. This is not surprising because similar conditions were used by Lloyd [14] to extend the Shannon–Nyquist theorem to a class of non band-limited processes. In this case the same condition, i.e., *disjointness property of the translates of the spectrum*, is useful to obtain rather simple expressions for the error formula. On the other hand, the general equation (1) suggests that the conditions used in, for example, [3], [6], and [17], may be sufficient but not necessary to have an error free prediction using uniform past samples.

In contrast, we will give *necessary and sufficient conditions* to predict a *not necessarily band-limited* w.s.s. process X_t, for all t, from $X_{-1}, X_{-2}\ldots$. These conditions naturally resemble the results of [14] and [1] where the reconstruction problem is solved using uniform samples from the past and future. Under this general conditions the prediction may become arbitrarily slow, however results such as [6] can be improved. We discuss this problem in Section II-D, Theorem 2.3. For this purpose, we recall that if X_t is a mean square continuous, wide sense stationary process there is a one to one correspondence between $\overline{\text{span}}\{X_t\}_{t\in\mathbb{R}}$ and $L^2(\mu) = \{f : \mathbb{R} \to \mathbb{C} / \int_{\mathbb{R}} |f|^2 d\mu < \infty\}$, where μ is the spectral measure of the process X_t, i.e., $R(t)$,

Manuscript received February 22, 2010; accepted September 19, 2010. Date of publication September 30, 2010; date of current version December 17, 2010. This work was partially supported by CONICET, the University of Buenos Aires grant No. I037, and ANPCYT Grant No. 209-2006. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Sofia C. Olhede.

The authors are with Dpto. Matemática, Facultad de Ingeniería, Universidad de Buenos Aires, Buenos Aires-CF 1063, Argentina, and Instituto Argentino de Matemática "A. Calderón"-CONICET (e-mail: jmedina@fi.uba.ar).

Digital Object Identifier 10.1109/TSP.2010.2081984

the correlation function of X_t, is the Fourier transform of the measure μ (Bochner). This fact will be used extensively. Also the following Kolmogorov theorem will be useful in our derivations. It shows that the absolutely continuous part of the measure is what really matters.

Theorem 1.1 [13, Ch. VII, p. 159]: Let m be a finite Borel measure on [0, 1], then

$$\inf_{P\in\mathcal{P}_0}\int_{[0,1]}|1-P|^2 dm = \inf_{P\in\mathcal{P}_0}\int_{[0,1]}|1-P|^2\frac{dm}{d\lambda}d\lambda$$

where $dm/d\lambda$ is the Radon–Nykodym derivative of the absolutely continuous part of m with respect to the Lebesgue measure λ and $\mathcal{P}_0$ denotes the space of polynomials in $e^{i2\pi x}$ of the form $P(e^{i2\pi x}) = \sum_{k=1}^{n} a_k e^{i2\pi kx}$. Note that $P : \partial U \longrightarrow \mathbb{C}$, where $\partial U = \{z \in \mathbb{C} : |z| = 1\}$. However, we shall treat each P as a function defined over the interval [0, 1), since there is a one to one correspondence between [0, 1) and ∂U.

The following classic result of Szegö [13, Ch. VII, p. 161] will be essential in our work.

Theorem 1.2: Let $w \geq 0$, $w \in L^1[0,1)$, then

$$\inf_{P\in\mathcal{P}_0}\int_{[0,1)}|1-P|^2 w d\lambda = \exp\left(\int_{[0,1)}\log(w)d\lambda\right)$$

where this infimum is understood to be equal to zero when $\log(w) \notin L^1[0,1)$.

II. Prediction Error and Disjointness Property of the Spectrum

A. Periodic Functions and Measures

We will see that some properties of uniform sampling can be derived from the properties of periodic functions and measures. For this purpose it is useful to consider the quotient space $\mathbb{R}/\mathbb{Z}$. We will denote the *canonical* projection $\pi : \mathbb{R} \longrightarrow \mathbb{R}/\mathbb{Z}$, the map that assigns to every $x \in \mathbb{R}$ its equivalence class $\pi(x)$. In our derivations, it is useful to make the following convention: to identify $\pi(x)$ with its *unique* representative in the interval [0, 1). That is to consider π as the following map: $\pi : \mathbb{R} \longrightarrow [0,1)$, $\pi(x) = \sum_{k\in\mathbb{Z}} \mathbf{1}_{[0,1)+k}(x)(x-k)$, where $\mathbf{1}_A$ is the indicator function of the set A.

Let $\mathbb{U} \subset \mathbb{R}$ be a Borel measurable set, then, the class of Borel subsets of $\mathbb{U}$ will be denoted by $\mathcal{B}(\mathbb{U})$. We recall that the induced measure π is the measure defined for every Borel set $U \subset [0,1)$ by the formula $\mu_\pi(U) = \mu(\pi^{-1}(U))$. By $\lambda(U)$, we mean the usual Lebesgue measure of U. We begin with a very intuitive result:

Proposition 2.1: Let μ be a Borel measure over $\mathbb{R}$ and let μ_π denote the measure over [0, 1) induced by the canonical projection to the quotient $\mathbb{R}/\mathbb{Z}$. If $(\mu_\pi)_s$ and $(\mu_\pi)_{\rm ac}$ denote the singular and absolutely continuous parts of μ_π, respectively, with respect to the Lebesgue measure, then:

a) for every Borel set U in [0, 1): $(\mu_\pi)_s(U) = \mu_s(\pi^{-1}(U))$ and $(\mu_\pi)_{\rm ac}(U) = \mu_{\rm ac}(\pi^{-1}(U))$; the measures μ_s and $\mu_{\rm ac}$ denote the singular and absolutely continuous parts of μ respectively, i.e., the singular part of the induced measure by π is the induced measure by π through the singular part of μ, the same for the absolutely continuous part;

b) if $w \in L^1(\mathbb{R}, d\lambda)$ is the Radon–Nykodym (R-N) derivative of $\mu_{\rm ac}$ with respect to Lebesgue measure, then $\sum_{k\in\mathbb{Z}} w(.+k)$ is the R-N derivative of $(\mu_\pi)_{\rm ac}$.

Proof: Part a)

We have the following measures defined for every Borel set U in [0, 1): $(\mu_s)_\pi(U) = \mu_s(\pi^{-1}(U))$ and $(\mu_{\rm ac})_\pi(U) = \mu_{\rm ac}(\pi^{-1}(U))$.

First, we will see that there exist $G, H \in \mathcal{B}[0,1)$ such that $G \cap H = \emptyset$, $[0,1) = G \cup H$, $(\mu_s)_\pi(G) = 0$ and $\lambda(H) = 0$. We know that there exist $H', G' \in \mathcal{B}(\mathbb{R})$ such that $G' \cup H' = \mathbb{R}$, $G' \cap H' = \emptyset$, $\mu_s(G') = 0$ and $\lambda(H') = 0$. Now, take $H = \pi(H')$ and $G = H^c = [0,1) \setminus H$, then

$$(\mu_s)_\pi(G) = \mu_s(\pi^{-1}(G)) = \mu_s(\pi^{-1}(H^c)) = \mu_s\left(\left(\pi^{-1}(H)\right)^c\right)$$

but $H' \subset \pi^{-1}(H)$, then

$$\mu_s\left(\left(\pi^{-1}(H)\right)^c\right) \leq \mu_s\left((H')^c\right) = \mu_s(G') = 0. \qquad (2)$$

On the other hand, let $I_k = [k, k+1)$, $k \in \mathbb{Z}$, then $H' = \bigcup_{k\in\mathbb{Z}} I_k \cap H'$, hence

$$H = \pi(H') = \bigcup_{k\in\mathbb{Z}} \left[(I_k \cap H') - k\right].$$

Since the Lebesgue measure is invariant under translations, then $\lambda[(I_k \cap H') - k] = \lambda(I_k \cap H') \leq \lambda(H') = 0$, and

$$\lambda(H) = \lambda\left(\bigcup_{k\in\mathbb{Z}}\left[(I_k \cap H') - k\right]\right) \leq \sum_{k\in\mathbb{Z}} \lambda\left[(I_k \cap H') - k\right] = 0$$

so that $(\mu_s)_\pi \perp \lambda$.

Now, we prove that $(\mu_{\rm ac})_\pi \ll \lambda$. Take $W \in \mathcal{B}[0,1)$ such that $\lambda(W) = 0$. Again, by the translation invariant property of Lebesgue measure: $\lambda(W+k) = \lambda(W) = 0$ so $\mu_{\rm ac}(W+k) = 0$ for every $k \in \mathbb{Z}$, since $\mu_{\rm ac} \ll \lambda$ over $\mathbb{R}$. Then

$$\begin{aligned}(\mu_{\rm ac})_\pi(W) &= \mu_{\rm ac}\left(\pi^{-1}(W)\right)\\ &= \mu_{\rm ac}\left(\bigcup_{k\in\mathbb{Z}} W+k\right) \leq \sum_{k\in\mathbb{Z}} \mu_{\rm ac}(W+k) = 0. \qquad (3)\end{aligned}$$

Finally, on the other hand, we have

$$\mu_\pi = (\mu_{\rm ac})_\pi + (\mu_s)_\pi. \qquad (4)$$

The equations given above, together with the uniqueness of the Lebesgue decomposition of a measure, show that (4) must be the Lebesgue decomposition of μ_π. The result of part a) follows from this.

Part b)

We will use periodization. Let $w = d\mu_{\text{ac}}/d\lambda$. Take $W \in \mathcal{B}[0,1)$, then, by (3)

$$(\mu_\pi)_{\text{ac}}(W) = (\mu_{\text{ac}})_\pi(W) = \int_{\bigcup_{k\in\mathbb{Z}} W+k} w d\lambda$$
$$= \sum_{k\in\mathbb{Z}} \int_{W+k} w(x)dx = \sum_{k\in\mathbb{Z}} \int_W w(x+k)dx$$
$$= \int_W \sum_{k\in\mathbb{Z}} w(x+k)dx.$$

■

B. Prediction Error

Recall that if μ is the spectral measure of a random process, then μ is a finite non negative Borel measure. Let us find an expression of the estimation error when we make a linear prediction using past samples taken at uniform rate. First, let us consider some general facts concerning periodic functions. Let f be a non negative Borel measurable 1-periodic function, i.e., $f(x) = f(x+1)$ for every $x \in \mathbb{R}$ (we will not distinguish between two functions which are equal for almost every $x[\mu]$). Hence, if we denote $f|_{[0,1)}$ the restriction of f to the interval [0, 1), i.e., $(f|_{[0,1)} \circ \pi)(x) = (f \circ \pi)(x) = f(\pi(x)) = f(x)$ for every $x \in \mathbb{R}$, then,

$$\int_{\mathbb{R}} f d\mu = \int_{\mathbb{R}} f|_{[0,1)} \circ \pi d\mu = \int_{[0,1)} f|_{[0,1)} d\mu_\pi = \int_{[0,1)} f d\mu_\pi.$$

For a non negative measurable f, this integrals are always well defined and all of them may diverge if one of them do so. This expression can be extended to arbitrary measurable f. By decomposing f in its non negative and negative parts, f^+ and f^- respectively, and provided that the integrals of both functions do not diverge simultaneously. In particular, this expression holds if we take $f = |1-P|^2$, with $P \in \mathcal{P}_0$. For this case:

$$\inf_{P\in\mathcal{P}_0} \int_{\mathbb{R}} |1-P|^2 d\mu = \inf_{P\in\mathcal{P}_0} \int_{[0,1)} |1-P|^2 d\mu_\pi. \quad (5)$$

To simplify, we will make the following simplification of the notation for the R-N derivatives of the absolutely continuous parts of the measures involved, calling: $d(\mu_\pi)_{\text{ac}}/d\lambda = d\mu_\pi/d\lambda$ and $d\mu_{\text{ac}}/d\lambda = d\mu/d\lambda$, etc.

The prediction error will depend on the following result:

Lemma 2.1: Let μ be a finite Borel measure and denote $A = \{x : (d\mu/d\lambda)(x) \neq 0\}$ then:

i)

$$\inf_{P\in\mathcal{P}_0} \int_{\mathbb{R}} |1-P|^2 d\mu = \begin{cases} \exp\left(\int_{[0,1)} \log\left(\frac{d\mu_\pi}{d\lambda}\right) d\lambda\right) & if\ \log\left(\frac{d\mu_\pi}{d\lambda}\right) \in L^1([0,1)) \\ 0 & \text{if } \lambda([0,1) \setminus \pi(A)) > 0 \\ 0 & \text{if } \log\left(\frac{d\mu_\pi}{d\lambda}\right) \notin L^1([0,1)) \end{cases}$$

ii) If $\lambda(A \cap A+k) = 0, \forall k \in \mathbb{Z}\setminus\{0\}$ then

$$\inf_{P\in\mathcal{P}_0} \int_{\mathbb{R}} |1-P|^2 d\mu = \begin{cases} \exp\left(\int_A \log\left(\frac{d\mu}{d\lambda}\right) d\lambda\right) & \text{if } \log\left(\frac{d\mu}{d\lambda}\right) \in L^1(A) \\ 0 & \text{if } \lambda([0,1) \setminus \pi(A)) > 0 \\ 0 & \text{if } \log\left(\frac{d\mu_\pi}{d\lambda}\right) \notin L^1(A) \end{cases}$$

iii) $\inf_{P\in\mathcal{P}_0} \int_{\mathbb{R}} |1-P|^2 d\mu = 0$ if and only if $\{e^{ik2\pi x}\}_{k\in\mathbb{N}}$ is complete in $\{f \in L^2(\mu) : f \text{ is } 1-\text{periodic}\}$.

The proof is rather straightforward. It will be clear that we can prove a) and b) simultaneously.

Remark: c) enables us to treat the problem of prediction as a completeness problem as in [6]. A similar idea is behind the proof of the sampling theorem of [14] using the one to one correspondence between the subspace generated by the samples $\ldots X_{-1}, X_0, X_1 \ldots$ and the subspace $\{f \in L^2(\mu) : f \text{ is } 1-\text{periodic}\}$.

Also note that the condition $\lambda(A \cap A+k) = 0, \forall k \in \mathbb{Z}\setminus\{0\}$ implies $\lambda(A) \leq 1 < \infty$.

Proof: Step 1: Let $w = d\mu_{\text{ac}}/d\lambda$, and $u = d\mu_{\pi ac}/d\lambda$). If we suppose that $\log(u) \in L^1([0,1))$, if $A = \{x \in \mathbb{R} : w(x) > 0\}$ in this case we have that $1 - \lambda(\pi(A)) = 0$. We can apply Szegö's theorem, together with Kolmogorov's Theorem 1.1 [13] to the right-hand side of (5) to get

$$\inf_{P\in\mathcal{P}_0} \int_{\mathbb{R}} |1-P|^2 d\mu = \exp\left(\int_{[0,1)} \log(u) d\lambda\right). \quad (6)$$

Step 2: If $\lambda(A \cap A+k) = 0, \forall k \in \mathbb{Z}\setminus\{0\}$, we prove that $\int_{\pi(A)} \log(u) d\lambda = \int_A \log(w) d\lambda$. For this, take $f(t) = t - \log(t) > 0$, for all $t > 0$. Recalling Proposition 2.1, we have that $\log(u(x)) = \sum_{k\in\mathbb{Z}} \log(w(x+k))$ and $f(u(x)) = \sum_{k\in\mathbb{Z}} f(w(x+k))$, for all $x \in \pi(A)$. Then since f is non negative:

$$\int_{\pi(A)} f(u) d\lambda = \sum_{k\in\mathbb{Z}} \int_{\pi(A)} f(w(.+k)) d\lambda \quad \text{(by monotone convergence).}$$

Now, since $\pi(A) = \bigcup_{k\in\mathbb{Z}} (A-k) \cap [0,1)$, and again from the disjointness condition on A, we have

$$\sum_{k\in\mathbb{Z}} \int_{\pi(A)} f(w(.+k)) d\lambda = \sum_{k\in\mathbb{Z}} \int_{[0,1)} f(w(.+k)) \mathbf{1}_A(.+k) d\lambda$$
$$= \sum_{k\in\mathbb{Z}} \int_{[k,k+1)} f(w) \mathbf{1}_A d\lambda$$
$$= \int_{\mathbb{R}} f(w) \mathbf{1}_A d\lambda$$

by the invariance property of the Lebesgue measure. Now, recalling that $w \in L^1(\mathbb{R})$ and Proposition 2.1—Part b: $\infty > \int_A w d\lambda = \int_{\pi(A)} u d\lambda$, then

$$\int_A \log(w) d\lambda = \int_{\pi(A)} \log(u) d\lambda. \tag{7}$$

These last integrals may be divergent, in this case they take the $-\infty$ value, since $w \in L^1(\mathbb{R})$. If we suppose that $\log(u) \in L^1([0,1))$, we have that $1 - \lambda(\pi(A)) = 0$. We can combine (6) with (7) to get

$$\inf_{P\in\mathcal{P}_0} \int_{\mathbb{R}} |1-P|^2 d\mu = \exp\left(\int_A \log(w) d\lambda\right). \tag{8}$$

Step 3:
Define $u_n = \max((1/n), u)$ and $d\nu_n = u_n d\lambda + d\mu_{\pi s}$. Then

$$\begin{aligned}\inf_{P\in\mathcal{P}_0} \int_{\mathbb{R}} |1-P|^2 d\mu &= \inf_{P\in\mathcal{P}_0} \int_{[0,1)} |1-P|^2 d\mu_\pi \\ &\le \inf_{P\in\mathcal{P}_0} \int_{[0,1)} |1-P|^2 d\nu_n = \exp\left(\int_{[0,1)} \log(u_n) d\lambda\right) \\ &= n^{-1+\lambda(\pi(A))} \exp\left(\int_{\pi(A)} \log(u_n) d\lambda\right). \end{aligned} \tag{9}$$

Equation (9) follows from Szegö's theorem and Kolmogorov's Theorem 1.1, since

$$\int_{[0,1)} \log(u_n) d\lambda = -\lambda([0,1) \setminus \pi(A)) \log(n) + \int_{\pi(A)} \log(u_n) d\lambda.$$

Step 4: On the other hand we have that $u_n(x) \ge u_{n+1}(x)$ for all $n \in \mathbb{N}$ and $x \in [0,1)$. It is also easy to verify that: $u_1 \in L^1(\pi(A))$, $\log(u_n) \ge \log(u_{n+1})$ and that $u_1 \ge u_n \ge \log(u_n)$ on $\pi(A)$, since $f(t) = t - \log(t) > 0$ for all $t > 0$. Then we have that $u_1 - \log(u_n) \underset{n\to\infty}{\nearrow} u_1 - \log(u)$ a.e. on $\pi(A)$. Then by the monotone convergence theorem:

$$\int_{\pi(A)} (u_1 - \log(u_n)) d\lambda \underset{n\to\infty}{\longrightarrow} \int_{\pi(A)} (u_1 - \log(u)) d\lambda.$$

But $u_1 \in L^1(\pi(A))$, then

$$\int_{\pi(A)} \log(u_n) d\lambda \underset{n\to\infty}{\longrightarrow} \int_{\pi(A)} \log(u) d\lambda. \tag{10}$$

Thus, if $\lambda(A \cap A + k) = 0, \forall k \in \mathbb{Z} \setminus \{0\}$ from (7) and (10):

$$\int_{\pi(A)} \log(u_n) d\lambda \underset{n\to\infty}{\longrightarrow} \int_A \log(w) d\lambda \tag{11}$$

where the right hand integral can be finite or $-\infty$. Taking $n \longrightarrow \infty$ in (9), all these results can be summarized as follows:

$$\inf_{P\in\mathcal{P}_0} \int_{\mathbb{R}} |1-P|^2 d\mu = \begin{cases} 0 & \text{if } \lambda([0,1) \setminus \pi(A)) > 0, \quad \text{by step 3 and (10)} \\ 0 & \text{if } \int_{\pi(A)} \log\left(\frac{d\mu_\pi}{d\lambda}\right) d\lambda = -\infty, \quad \text{by step 4 and (10)} \\ 0 & \text{if } \lambda(A \cap A + k) = 0 \forall\, k \in \mathbb{Z} \\ & \text{and } \int_A \log\left(\frac{d\mu_\pi}{d\lambda}\right) d\lambda = -\infty, \quad \text{by step 4 and (11).} \end{cases}$$

c) The condition $\inf_{P\in\mathcal{P}_0} \int_{\mathbb{R}} |1-P|^2 d\mu = 0$ is necessary. For the sufficiency, note that $\{e^{ik2\pi x}\}_{k\in\mathbb{Z}}$ is always complete in $\{f \in L^2(\mu) : f \text{ is } 1-\text{periodic}\}$, so the result follows if we prove that $e^{i2\pi n x} \in \overline{\text{span}}\{e^{ik2\pi x}\}_{k\in\mathbb{N}}$ for every $n = 0, -1, -2, \ldots$. For this purpose note that

$$\begin{aligned}\int_{[0,1)} \left| e^{-i2\pi x} - \sum_{k=1}^{N} a_k e^{ik2\pi x} \right|^2 d\mu_\pi(x) \\ = \int_{[0,1)} \left| (1-a_1) - \sum_{k=1}^{N-1} a_{k+1} e^{ik2\pi x} \right|^2 d\mu_\pi(x).\end{aligned}$$

Then

$$\inf_{P\in\mathcal{P}_0} \int_{[0,1)} \left| e^{-i2\pi x} - P(e^{i2\pi x}) \right|^2 \mu_\pi(x) = 0,$$

and from this, we get the result inductively. ∎

As a corollary of the previous result we obtain,

Corollary 2.1: Let X_t be a w.s.s. process with spectral measure μ. Suppose that $\exists A \in \mathcal{B}(\mathbb{R})$ such that $\mu_{ac}(A^c) = 0$. Then, for fixed $t \in \mathbb{R}$:

a) if $\lambda([0,1) \setminus \cup_{k\in\mathbb{Z}} A + k) > 0$ then:

$$\inf_{a_k\in\mathbb{C},\, n\in\mathbb{N}} \mathbb{E} \left| X_t - \sum_{k=1}^{n} a_k X_{t-k} \right|^2 = 0; \tag{12}$$

b) if $\lambda([0,1) \setminus \cup_{k\in\mathbb{Z}} A + k) = 0$ and $\mu_{ac}(A \cap A + k) = 0$, $k \in \mathbb{Z} \setminus 0$, then

$$\inf_{a_k\in\mathbb{C},\, n\in\mathbb{N}} \mathbb{E} \left| X_t - \sum_{k=1}^{n} a_k X_{t-k} \right|^2 = \exp\left(\int_A \log\left(\frac{d\mu}{d\lambda}\right) d\lambda\right). \tag{13}$$

Remark: Note that the condition $\mu_{ac}(A \cap A + k) = 0$ is not necessary for a). See Example 4. On the other hand, this condition is also equivalent to the existence of a measurable subset B such that $\lambda(B \cap B + k) = 0$ for all $k \in \mathbb{Z} \setminus \{0\}$, and $\mu_{ac}(B^c) = 0$.

Proof: Fix $t \in \mathbb{R}$. Denote $\hat{X}_t = \sum_{k=1}^{n} a_k X_{t-k}$. Then, the result is immediate from Lemma 2.1, since

$$\mathbb{E}|X_t - \hat{X}_t|^2 = \int_A \left| e^{i2\pi xt} - \sum_{k=1}^{n} a_k e^{i2\pi x(t-k)} \right|^2 d\mu(x)$$
$$= \int_A \left| 1 - \sum_{k=1}^{n} a_k^* e^{i2\pi kx} \right|^2 d\mu(x).$$

■

Remark: In particular, when $\int_A \log((d\mu/d\lambda)(x))dx = -\infty$, we define that (13) equals 0. We recall that a similar disjointness condition of the spectrum was successfully used in [1] and [14] to extend the classical Shannon–Nyquist–Kotelnikov sampling theorem to the non band-limited case. On the other hand, if A is an interval, we recover the result of [1] for band-limited processes. In particular, this also implies the results of [6], [17], since for a band-limited signal condition a) is equivalent to sampling faster than the Nyquist sampling rate.

C. Necessary and Sufficient Conditions for a Process to Be Predictable Using Uniform Past Samples

Lemma 2.1 gives an expression for the prediction error. Now, we will prove that the disjointness property of the spectrum together with the divergence of a logarithmic integral of the spectral density are necessary and sufficient conditions for a (w.s.s.) process to be predictable for all t using uniform past samples. For this purpose we recall:

Theorem 2.1 [14]: Let X_t, $t \in \mathbb{R}$ be a w.s.s. process with spectral measure μ, the following are equivalent:

i) $\overline{\text{span}}\{X_k\}_{k\in\mathbb{Z}} = \overline{\text{span}}\{X_t\}_{t\in\mathbb{R}}$, (with respect to the $L^2(\Omega, \mathcal{F}, \mathbb{P})$ norm);

ii) $\overline{\text{span}}\{e^{i2\pi kx}\}_{k\in\mathbb{Z}} = \{f \in L^2(\mu) : f \; is \; 1-$periodic$\} = L^2(\mu)$. In particular, this means that $L^2(\mu)$ consists of all f such that $f(x+1) = f(x)$ for almost all x $[\mu]$, and $\int_{\mathbb{R}} |f|^2 d\mu < \infty$;

iii) there exists $A \in \mathcal{B}(\mathbb{R})$ such that $\mu(A^c) = 0$ and $A \cap A + k = \oslash$, $k \in \mathbb{Z} \setminus \{0\}$.

Now, we can completely characterize (w.s.s.) processes which are completely predictable from uniform past samples:

Theorem 2.2: Let X_t, $t \in \mathbb{R}$ be a w.s.s. process with spectral measure μ, the following are equivalent:

i) there exists $A \in \mathcal{B}(\mathbb{R})$ such that $\mu(A^c) = 0$ and $A \cap A + k = \oslash$, $k \in \mathbb{Z} \setminus \{0\}$; and: $\int_A \log(d\mu/d\lambda)d\lambda = -\infty$ or $\lambda([0,1) \setminus \pi(A)) > 0$;

ii) for all $t \in \mathbb{R}$:

$$\inf_{a_k \in \mathbb{C},\, n\in\mathbb{N}} \mathbb{E}\left| X_t - \sum_{k=1}^{n} a_k X_{-k} \right|^2 = 0$$

i.e., $\overline{\text{span}}\{X_k\}_{k\in\mathbb{N}} = \overline{\text{span}}\{X_t\}_{t\in\mathbb{R}}$, (with respect to the $L^2(\Omega, \mathcal{F}, \mathbb{P})$ norm).

Remark: Note that condition i) of Theorem 2.2 is slightly different of (1). In contrast to (1) the integration is made only over the support of the measure and the weight $(1+x^2)^{-1}$ is missing. Equation (1) is a weaker condition. It may be possible that $\overline{\text{span}}\{X_k\}_{k\in\mathbb{N}} \neq \overline{\text{span}}\{X_t\}_{t\in\mathbb{R}}$ and that simultaneously, for $s < \infty$, $\overline{\text{span}}\{X_t\}_{t\leq s} = \overline{\text{span}}\{X_t\}_{t\in\mathbb{R}}$.

Proof:

i) ⇒ ii) Is equivalent to prove that

$$\overline{\text{span}}\{e^{i2\pi kx}\}_{k\in\mathbb{N}} = L^2(\mu).$$

Under these conditions by Lemma 2.1 (b): $\inf_{P\in\mathcal{P}_0} \int_{\mathbb{R}} |1 - P|^2 d\mu = 0$, then, first, by Lemma 2.1 (c) and then by Theorem 2.1: $\overline{\text{span}}\{e^{i2\pi kx}\}_{k\in\mathbb{N}} = \{f \in L^2(\mu) : f \; is \; 1 - periodic\} = L^2(\mu)$.

ii) ⇒ i) In particular ii) implies $\inf_{a_k\in\mathbb{C}, n\in\mathbb{N}} \mathbb{E}|X_j - \sum_{k=1}^{n} a_k X_{-k}|^2 = 0$, for $j = 0, 1, 2, \ldots$. It is equivalent,

$$\overline{\text{span}}\{X_k\}_{k\in\mathbb{Z}} = \overline{\text{span}}\{X_k\}_{k\in\mathbb{N}} \tag{14}$$

but this also implies $\inf_{P\in\mathcal{P}_0} \int_{\mathbb{R}} |1 - P|^2 d\mu = 0$. On the other hand, ii) together with (14) is equivalent to $\overline{\text{span}}\{X_k\}_{k\in\mathbb{N}} = \overline{\text{span}}\{X_t\}_{t\in\mathbb{R}}$. Then, by Theorem 2.1, there exists $A \in \mathcal{B}(\mathbb{R})$ such that $\mu(A^c) = 0$ and $A \cap A + k = \oslash$, $k \in \mathbb{Z} \setminus \{0\}$. And by Lemma 2.1 $\int_A \log(d\mu/d\lambda)d\lambda = -\infty$ or $\lambda([0,1) \setminus \pi(A)) > 0$. ■

Remark: In particular, Theorem 2.2 implies that if $\{X_t\}_{t\in\mathbb{R}}$ is band-limited, not only we can predict X_t from the past samples $X_{t-1}, X_{t-2}, \cdots$ as claimed in [17] and [6] but also $X_{t+\tau}$, for all $\tau \in \mathbb{R}$.

D. Some Examples and Some Remarks

In this section, we give some simple examples, and we discuss the problem of the convergence rate.

Example 1. (Spectral Density With Unbounded Support and Finite Measure): The processes we have considered are not necessarily finite sums of band-limited processes. For example, consider the intervals $I_k = [1 - 2^{-k}, 1 - 2^{-k-1})$, $k \in \mathbb{N}$ and define a measure μ by:

$$\frac{d\mu_\pi}{d\lambda}(x) = \sum_{k=1}^{\infty} \mathbf{1}_{I_k + k}(x).$$

Then there exists a w.s.s. process, say X_t, with spectral density $d\mu/d\lambda$ defined as above, and $d\mu/d\lambda$ verifies that $(d\mu_\pi/d\lambda)(x) = \mathbf{1}_{[(1/2),1)}(x)$ for every $x \in [0,1)$. Then, this process verifies the conditions of Theorem 2.2 for X_t to be predictable: i.e., for all t: X_t is linearly determined by the samples $X_{-1}, X_{-2}, \ldots$.

Example 2. (Sampling at Nyquist Rate): We note that some band-limited processes can be predicted with samples taken at *exactly* the Nyquist rate. For example, we can build a process with spectral density $(d\mu/d\lambda)(x) = exp(-(1/|x|))\mathbf{1}_{[-(1/2),(1/2)]}(x)$. This process conforms to the conditions of Theorem 2.2, and again, for all t, X_t is linearly determined by the samples $X_{-1}, X_{-2}, \ldots$.

In the following example, for a *fixed* $t \in \mathbb{R}$, we want to predict from samples $X_{t-1}, X_{t-2}, \ldots$:

Example 3. (Uniform Sampling With a Spectral Density Which Has Full Measure): Although the formulas of part a) of Lemma 2.1 are not as simple as those of part b), they give the more general condition for a process to be predictable at a fixed time t from samples taken at uniform rate. For example consider a process with spectral measure defined by

$$\frac{d\mu_\pi}{d\lambda}(x) = \sum_{n\in\mathbb{Z}} 2^{-|n|}\mathbf{1}_{[n,n+1)}(x) exp\left(-\frac{1}{|x-n|}\right).$$

Then, is easy to check that

$$\frac{d\mu_\pi}{d\lambda}p(x) = 3\, exp\left(-\frac{1}{|x|}\right)\mathbf{1}_{[0,1)}(x).$$

Example 4. (A Case Where It is Possible to Find Coefficients Independent of μ): As we have seen, in general there is no explicit formula for the predictor's coefficients. However we notice that imposing a boundedness condition on the power spectrum, in some cases, such as when $\lambda([0,1)\setminus\pi(A)) > 0$ in Theorem 2.2 (for band-limited processes this is equivalent to taking samples at a rate over Nyquist's rate), it is possible to find coefficients independent of μ. From the orthogonality principle, with a similar argument to [5] we have that for fixed $t\in\mathbb{R}$ we can find $a_{kn}(t)$ (generally non optimal) such that $\mathbb{E}|X_t - \sum_{k=1}^n a_{kn}(t)X_{-k}|^2 = e(n)\to 0$ as $n\to\infty$. These coefficients are determined by the equations: $\sum_{k=1}^n a_{kn}(t)\mathbf{1}_A^\vee(k-j) = \mathbf{1}_A^\vee(t-j)$ for $j=1\ldots n$. Example 1 is a case of this fact. This naturally resembles a result from [1]. So despite there is no explicit formula, under some additional conditions the coefficients are determined implicitly and are independent of μ.

1) Rate of Convergence: Theorem 2.2 gives necessary and sufficient conditions for a process to be predictable from its past samples. Under these general conditions convergence could become arbitrarily slow. However the importance of this result is that it gives a closed answer to the question of how much results such as [5] and [7] can be improved. Let us give an example. The following result, contains [5] and [7] as particular cases and shows that we still can have exponential rate of convergence when the process is not band-limited and not sampled above twice the Nyquist rate as in [7].

Theorem 2.3:

a) Let $w\in L^1(\mathbb{R})$ be non-negative and $A=\{x: w(x)\neq 0\}$ be such that: $\lambda(A\cap A+k)=0$, for all $k\in\mathbb{Z}\setminus\{0\}$, and $\overline{\pi(A)}\subsetneq[0,1)$ then, given $0<\beta<1$ there exists a sequence of polynomials in $e^{-i2\pi x}$ with zero independent term $\{P_n(e^{-i2\pi x})\}_n$ such that

$$\int_{\mathbb{R}} |1-P_n|^2 w\,d\lambda \le \|w\|_{L^1(\mathbb{R})}\beta^n.$$

b) In particular if $\{X_t\}_{t\in\mathbb{R}}$ is a w.s.s. process with an spectral density w that verifies the conditions of a), then there exists a sequence of predictors $\{\hat{X}_t^n\}_n$ which converges to X_t in $L^2(\Omega,\mathcal{F},\mathbb{P})$ with exponential rate and a.s.

To prove this we need the following result from the theory of Harmonic functions:

Theorem 2.4: [12] Let E be proper closed subset of [0, 1). Any continuous function on E can be approximated uniformly by Taylor polynomials $\sum_{n=0}^N a_n e^{in2\pi x}$.

Important Remark: This result *is not* the Weierstrass approximation theorem, nor a consequence of the completeness theorem of the trigonometric system in $L^2[0,1)$ either. See for example p. 115 of [12].

Proof of Theorem 2.3 (Part a): Given $Q(e^{i2\pi x}) = \sum_{k=0}^m a_k e^{ik2\pi x}$. Then $|e^{-i2\pi x} - Q(e^{i2\pi x})| = |1 - \sum_{k=0}^m a_k^* e^{-i(k+1)2\pi x}|$ for all x. Now given $0<\beta<1$, taking $\delta=\beta^{1/2}$, by Theorem 2.4 there exists $Q_0(e^{i2\pi x}) = \sum_{k=0}^{m_0} b_k e^{ik2\pi x}$ such that $\sup_{x\in\overline{\pi(A)}} |e^{-i2\pi x} - Q_0(e^{i2\pi x})| < \delta$. But from the previous argument this is equivalent to

$$\sup_{x\in\overline{\pi(A)}}\left|1-\sum_{k=0}^{m_0} b_k^* e^{-i(k+1)2\pi x}\right| < \delta. \tag{15}$$

Write $R(e^{i2\pi x}) = \sum_{k=0}^{m_0} b_k^* e^{-i(k+1)2\pi x}$, then $(1-R)^n = 1+\sum_{j=0}^{n-1}\binom{n}{j}(-R)^{n-j}$. So choosing

$$P_n = \sum_{j=0}^{n-1}\binom{n}{j}(-1)^{n-j+1}(R)^{n-j},$$

by (15) we have

$$\begin{aligned}\int_{\mathbb{R}} |1-P_n|^2 w\,d\lambda &= \int_{[0,1)} |1-P_n|^2 w_\pi\,d\lambda\\ &= \int_{\pi(A)} |1-P_n|^2 w\,d\lambda \le \beta^n \int_{[0,1)} w_\pi\,d\lambda.\end{aligned}$$

(*Part b*) The first part is immediate from a) and the a.s. convergence follows from the Borel–Cantelli lemma: pick $\epsilon>0$ then by Tchevicheff's inequality,

$$\mathbb{P}\left(\left|X_t-\hat{X}_t^n\right|>\epsilon\right) \le \frac{\mathbb{E}\left|X_t-\hat{X}_t^n\right|^2}{\epsilon^2} \le \frac{\beta^n}{\epsilon^2}$$

so $\sum_n \mathbb{P}(|X_t-\hat{X}_t^n|>\epsilon)<\infty$ and then for $\epsilon>0$: $\mathbb{P}(\lim_{n\to\infty}\{|X_t-\hat{X}_t^n|>\epsilon\})=0$ and the assertion follows from this. ■

2) Remark: Note that to build the sequence $\{P_n\}_n$ we need an initial polynomial Q satisfying (15) and this determines the coefficients of the whole sequence of predictors. In some cases this initial polynomial can be found by a direct method as in [7, Lemma of Sec. II, p. 411], however in a more general setting, noting that $\{1, e^{i2\pi kx}\}_{k=1\ldots m}$ satisfies the "Haar condition" [9], one could use some approximation algorithm in uniform norm to find it. Finally, we have not lost generality in this proof assuming that the spectral measure μ is absolutely continuous since we could appeal to Theorem 1.1 or to the previous results to treat such a case.

E. Non Uniform Samples

We will briefly discuss the problem of prediction using non uniform samples. The problem of prediction using uniform past samples is equivalent to show that the set $\{e^{i2\pi kx}\}_{k\in\mathbb{N}}$ is complete in $L^2(\mu)$. To obtain an analogous result for non uniform samples taken at times $0>t_1>t_2>\ldots$, one way is to give sufficient conditions for the completeness of $\{e^{it_k x}\}_{k\in\mathbb{N}}$. Different conditions can be used, mainly involving techniques from non harmonic Fourier series [16], [21]. Generally these results depend on the properties of the zeroes of analytic functions and related results such as Carleman's theorem [4]. Among the several conditions that can be given we state a brief result, which resembles the Müntz–Szász theorem [16], [20]. On the other hand, this result again relates the prediction problem with the condition of the divergence of a logarithmic integral of the spectral density as in the uniform samples case. Let us recall the following result from complex analysis:

Theorem 2.5 [20, Chap. 15]: Let $F(z)$ be holomorphic and bounded over $U=\{z\in\mathbb{C}:|z|<1\}$, let $\{z_1,z_2,\ldots\}\subset U$ be the set of zeros of F, if

$$\sum_{n\in\mathbb{N}}(1-|z_n|)=\infty$$

then $F\equiv 0$.

With this in mind, we can prove:

Theorem 2.6: Let $0<t_0<t_1<\ldots$ be a sequence such that $t_n \underset{n\to\infty}{\longrightarrow} \infty$ and

$$\sum_{n\in\mathbb{N}}\left(1-\frac{t_n}{(4+t_n^2)^{\frac{1}{2}}}\right)=\infty,$$

then if μ is a finite Borel measure supported over a measurable set $A\subseteq[0,1]$, such that

$$\int_A \log\left(\frac{d\mu}{d\lambda}\right)d\lambda=-\infty, \tag{16}$$

then $\{e^{it_nx}\}_{n\in\mathbb{N}}$ is complete in $L^2(\mu)$.

Proof: It suffices to prove that given $f\in L^2(\mu)$, if $\int_{\mathbb{R}} e^{i2\pi\lambda t_n}f(\lambda)d\mu(\lambda)=0,\ n=0,1\ldots$ then $f\equiv 0$. Define the function on the complex variable $z=x+iy$ over the domain $D=\{z=x+iy: y>-1\}$, by

$$F(z)=\int_{\mathbb{R}} e^{i2\pi\lambda z}f(\lambda)d\mu(\lambda),$$

then $F(t_n)=0,\ n=0,1,\ldots$. Since the measure μ is supported over the interval [0, 1], then by the Cauchy–Schwartz inequality

$$\begin{aligned}|F(z)|&\le\left(\int_{\mathbb{R}} e^{-4\pi\lambda y}d\mu(\lambda)\right)^{\frac{1}{2}}\|f\|_{L^2(\mu)}\\&\le(\mu([0,1]))^{\frac{1}{2}}\|f\|_{L^2(\mu)}e^{2\pi}\quad\forall z\in D.\end{aligned}$$

So, F is a bounded function over D. Moreover, in a similar manner is easy to check that F is continuous. By Morera's theorem F is holomorphic over D: take a closed triangle $\triangle\subset D$ then

$$\begin{aligned}\oint_{\partial\triangle}F(z)dz&=\oint_{\partial\triangle}\left\{\int_{\mathbb{R}} e^{i2\pi\lambda z}f(\lambda)d\mu(\lambda)\right\}dz\\&=\int_{\mathbb{R}}\left\{\oint_{\partial\triangle}e^{i\lambda z}dzf(\lambda)\right\}d\mu(\lambda)=0.\end{aligned}$$

The previous calculations show that we can apply Fubini's theorem to interchange the order of integration. Since the inner integral is zero by Cauchy's theorem, then F is holomorphic over D. Now let us consider the conformal mapping:

$$\psi(z)=i\left(\frac{1+z}{1-z}\right)-i.$$

The mapping ψ is a one to one mapping from $U=\{z:|z|<1\}$ to D. Now define $G:U\longrightarrow\mathbb{C}$ as $G(z)=(F\circ\psi)(z)$, then G has zeros: $z_n=\psi^{-1}(t_n)$. This implies

$$\sum_{n\in\mathbb{N}}(1-|z_n|)=\sum_{n\in\mathbb{N}}\left(1-\frac{t_n}{(4+t_n^2)^{\frac{1}{2}}}\right)=\infty,$$

then by Theorem 2.5 $G\equiv 0$, or equivalently $F\equiv 0$, so in particular $F(n)=0,\ n\in\mathbb{N}$, but under the condition given by (16) this implies $f\equiv 0$. ■

From this, one gets immediately:

Corollary 2.2: Let X_t be a w.s.s. process with spectral measure μ supported over a measurable set $A\subseteq[0,1]$, such that

$$\int_A log\left(\frac{d\mu}{d\lambda}\right)d\lambda=-\infty. \tag{17}$$

Then, given a set of samples $\{X_{t_k}\}_{k\in\mathbb{N}}$, with $\ldots t_2<t_1<0$ a decreasing sequence, such that

$$\sum_{n\in\mathbb{N}}\left(1-\frac{|t_n|}{(4+t_n^2)^{\frac{1}{2}}}\right)=\infty, \tag{18}$$

then, for all $t\in\mathbb{R}$:

$$\inf_{a_k\in\mathbb{C},\ n\in\mathbb{N}}\left(\mathbb{E}\left|X_t-\sum_{k=1}^{n}a_kX_{t_k}\right|^2\right)^{\frac{1}{2}}=0.$$

Example: An example of non uniform samples that verify the condition given by (18), is to take $t_n=-(2n-((6n-4)/(2n-1)))^{1/2}$.

III. Conclusion

We gave conditions for the prediction of a w.s.s. random process at an instant t, from past samples $X_{t-1},X_{t-2},\cdots$. Also we gave a general expression for the prediction error when the exact prediction conditions are not satisfied. Moreover, we gave necessary and sufficient conditions for X_t to be predicted with arbitrarily small error for all $t\in\mathbb{R}$, from past samples $X_{-1},X_{-2},\cdots$, see Theorem 2.2. This result is equivalent to first predict $X_0,X_1,\cdots$, and then interpolate. This is the more general form in which the problem stated in previous work, such as [3], [6], and [17], can be formulated. We gave a very general condition in which error free prediction for not necessarily band-limited stationary processes is theoretically possible from uniform past samples. This is characterized by some properties of the translates of the support of the spectral measure, as in previous work [1], [2], and [14]. This is useful to limit the search of other more practical sufficient conditions such as those of the main theorem in [7]. Moreover since our proofs rely on periodization techniques, "universal" coefficients can be found under more general conditions than imposing the process to be band-limited, as in the case in which samples are taken at an appropriate rate as in [7] but where the spectrum verifies a more general condition such as in i) of our Theorem 2.2 or the main results of [1] and [14].

ACKNOWLEDGMENT

The authors would to like to thank the anonymous reviewers for their very helpful comments.

REFERENCES

[1] M. G. Beaty and M. M. Dodson, "The WKS theorem, spectral translates and Plancherel's formula," *J. Fourier Anal. Appl.*, vol. 10, no. 2, pp. 179–199, 2004.
[2] M. G. Beaty and J. R. Higgins, "Aliasing and poisson summation in the sampling theory of Paley-Wiener spaces," *J. Fourier Anal. Appl.*, vol. 1, no. 1, pp. 67–85, 1994.
[3] F. J. Beutler, "Sampling theorems and bases in a Hilbert space," *Inf. Control*, vol. 4, pp. 97–117, 1961.
[4] F. J. Beutler, "Error-free recovery of signals from irregularly spaced samples," *SIAM Rev.*, vol. 8, no. 3, pp. 328–335, 1966.
[5] J. L. Brown, "Uniform linear prediction of bandlimited processes from past samples," *IEEE Trans. Inf. Theory*, vol. 28, no. 4, pp. 662–664, 1972.
[6] J. L. Brown, "On the prediction of a band-limited signal from past samples," *Proc. IEEE*, vol. 74, no. 11, pp. 1596–1698, Nov. 1986.
[7] J. L. Brown and O. Morean, "Robust prediction of band-limited signals from past samples," *IEEE Trans. Inf. Theory.*, vol. 32, no. 3, pp. 410–412, 1986.
[8] S. Cambanis and E. Masry, "Truncation error bounds for the cardinal sampling expansions of band-limited signals," *IEEE Trans. Inf. Theory*, vol. 28, no. 4, pp. 605–612, 1982.
[9] E. W. Cheney, *Introduction to Approximation Theory*. New York: Chelsea, 1982.
[10] I. I. Gikhman and A. V. Skorokhod, *The Theory of Stochastic Processes*. New York: Springer, 1974, vol. I.
[11] U. Grenander and M. Rosenblatt, *Statistical Analysis of Stationary Time Series*. New York: Wiley, 1957.
[12] Y. Katznelson, *An Introduction to Harmonic Analysis*, 3rd. ed. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[13] P. Koosis, *Introduction to H_p Spaces*. Cambridge, U.K.: Cambridge Univ. Press, 1998.
[14] S. P. Lloyd, "A sampling theorem for stationary (wide sense) stochastic processes," *Trans. Amer. Math. Soc.*, vol. 92, no. 1, pp. 1–12, 1959.
[15] J. M. Medina and B. Cernuschi-Frías, "On the prediction of a multiband signal from its past samples," in *Proc. 14th IEEE Statistical Signal Processing Workshop 2007*, pp. 220–223.
[16] R. E. A. C. Paley and N. Wiener, "Fourier transforms in the complex domain," *Amer. Math. Soc. Colloquium Publications*, vol. XIX, 1934.
[17] A. Papoulis, "A note on the predictability of band-limited processes," *Proc. IEEE*, vol. 73, no. 8, p. 1332, Aug. 1985.
[18] M. Rosenblatt, *Stationary Sequences and Random Fields*. Cambridge, MA: Birkhäuser Verlag, 1985.
[19] Y. A. Rozanov, *Stationary Random Processes*. San Francisco, CA: Holden-Day, 1967.
[20] W. Rudin, *Real and Complex Analysis*, 3rd. ed. New York: McGraw-Hill, 1986.
[21] R. M. Young, *An Introduction to Nonharmonic Fourier Series*. New York: Academic, 1980.
[22] L. A. Wainstein and V. D. Zubakov, *Extraction of Signals in Noise*. Englewood Cliffs, NJ: Prentice-Hall, 1962.

Juan Miguel Medina was born in Buenos Aires, Argentina, in 1977. He received the Licenciado en Ciencias Matemáticas degree in pure mathematics from the Universidad de Buenos Aires, Buenos Aires, Argentina. He is currently working towards the Ph.D. degree in mathematics from the Universidad de Buenos Aires, at the Insituto Argentino de Matemática "A. Calderón"-CONICET.

He is currently with the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).

Bruno Cernuschi-Frías (S'77–M'78–SM'01) was born in 1952, Montevideo, Uruguay. He received the Ph.D. degree in electrical engineering from Brown University, Providence, RI, in 1984.

He is currently a full tenured Professor with the Department of Electronics, Faculty of Engineering, University of Buenos Aires, Argentina, and Principal Researcher at the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).