SIGNAL SAMPLING AND RECOVERY UNDER LONG-RANGE DEPENDENT NOISE WITH APPLICATION TO LACK-OF-FIT TESTS

M. Pawlak

Electrical and Computer Engineering,
University of Manitoba,
Winnipeg, Manitoba, R3T5V6, Canada.
email : pawlak@ee.umanitoba.ca

U. Stadtmüller

Inst. für Mathematik,
University of Ulm,
89069 Ulm, Germany.
email:stamue@mathematik.uni-ulm.de

ABSTRACT

The paper examines the impact of the additive correlated noise on the accuracy of a signal reconstruction algorithm originating from the Whittaker-Shannon sampling interpolation formula. The proposed reconstruction method is a smooth post-filtering correction of the classical Whittaker-Shannon interpolation series. We assess the global accuracy of the proposed reconstruction algorithm for long memory stationary errors being independent on the sampling rate. We also examine a class of long memory noise processes for which the correlation function depends on the sampling rate. Exact rates at which the reconstruction error tends to zero are evaluated. We apply our theory to the problem of designing non-parametric lack-of-fit tests for verifying a parametric assumption on a signal. The theory of the asymptotic behavior of quadratic forms of stationary sequences is utilized in this case.

1. INTRODUCTION

The Whittaker-Shannon (WS) interpolation series plays a fundamental role in representing signals/images in the discrete domain. In fact, it is commonly recognized as a milestone in signal processing, communication systems, as well as Fourier analysis [1] . The WS reconstruction theorem says that if an analog signal $f(t)$ is band-limited with the bandwidth Ω (this in the sequel we shall denote as $f \in BL(\Omega)$) then it can be perfectly reconstructed from its discrete values $\{f(k\tau)\}$ by

$$f(t) = \sum_{k=-\infty}^{\infty} f(k\tau) \operatorname{sinc}(\pi\tau^{-1}(t - k\tau)), \qquad (1)$$

provided that $\tau \leq \pi/\Omega$, where $\operatorname{sinc}(t) = \sin(t)/t$. The WS interpolation series has been extended to a number of circumstances including multiple dimensions, random signals, not necessarily band-limited signals, sampling in generalized spaces, and reconstruction from irregularly sampled data [1] . In practice one rarely has access to perfect samples $\{f(k\tau)\}$ but rather to their noisy version $\{y_k\}$ due to measurement and transmission errors. This important issue has been mentioned often in the signal processing literature but no rigorous algorithms with established convergence properties for a signal reconstruction from sampled and noisy data were given. A naive approach would use the reconstruction method based on (1) with $\{f(k\tau)\}$ replaced by the samples $\{y_k\}$. This approach cannot work since it is dangerous to interpolate the noisy data. The first thorough treatment of this problem was given in [2], see also [3] for an overview of this problem. In all these contributions the white noise case has been mostly examined. In [4], however, the extension of the previous theory to short memory noise processes was made. The problem addressed in this paper is to provide the further generalization to the case of long-range dependent noise processes. There are many physical and man-made phenomena that exhibit strong long-term correlations [5], [6]. In many applications we are faced with the aggregation effect of random environments yielding a noise process which is not only correlated but also exhibits long-range dependence [5], [6]. Particularly, this takes place in wireless networks where the multipath random environment yields fading effect resulting in dependent noise with long memory effects. In this paper we consider the following statistical model. We observe $2n + 1$ noisy data points

$$y_k = f(k\tau) + \varepsilon_k, \quad |k| \leq n, \qquad (2)$$

and we wish to design a consistent reconstruction scheme resembling (1). Here $\{\varepsilon_k\}$ is the zero mean finite variance stationary stochastic process. The following model on the correlation structure of $\{\varepsilon_k\}$ is employed in this paper.

Assumption 1 *Let $\{\varepsilon_k\}$ be a weakly stationary stochastic process with $E\varepsilon_k = 0$, $\operatorname{var}(\varepsilon_t) = \sigma^2$, and $\operatorname{cov}(\varepsilon_{k+\ell}, \varepsilon_\ell) = r(|k|)$, such t hat*

$$r(k) = L(|k|)|k|^{-\alpha}, \quad 0 < \alpha \leq 1, \quad |k| \neq 0, \qquad (3)$$

where $L(\bullet)$ is a slowly varying function at ∞. The processes satisfying (3) is said to have long-range dependence (LDR) since $\sum_{k=1}^{\infty}|r(k)| = \infty$. On the contrary a process for which $\sum_{k=1}^{\infty}|r(k)| < \infty$ exhibits short-range dependence (SRD). It is worth noting that two popular *LRD* processes, namely fractional Gaussian noise and *FARIMA* processes meet Assumption 1.

The noise process satisfying Assumption 1 is independent on the sampling rate of the transmitted signal, i.e., the correlation function $r(k)$ of $\{\varepsilon_k\}$ is independent on τ. On the other hand the model in (2) can be viewed as the sampling process of the analog signal $f(t)$ submerged in a continuous time colored process $X(t)$ such that $\varepsilon_k = X(k\tau)$. In this case the correlation function $r(k)$ of $\{\varepsilon_k\}$ must depend on τ in such a way that $\{\varepsilon_k\}$ gets more dependent with higher oversampling, i.e., when τ gets smaller. This situation may appear during the actual sampling process of a signal imbedded in a continuous stationary stochastic process. This motivates the following model of dependence of $r(k)$ on τ.

Assumption 2 *Let $\{\varepsilon_k\}$ be a weakly stationary stochastic process with $E\varepsilon_k = 0$ and $E\{\varepsilon_k\varepsilon_{k+j}\} = r(\rho j\tau)$. Let moreover ρ depend on τ in such a way that*

$$\tau\rho_\tau \to \theta \quad as \quad \tau \to 0, \tag{4}$$

where $0 \leq \theta \leq \infty$.

The factor ρ_τ in (4) is a measure of the strength of the dependence of the noise process. In fact, if ρ_τ increases fast enough to infinity, e.g., $\rho_\tau \simeq d\tau^{-\beta}$, $\beta > 1$, as $\tau \to 0$ then the noise process is almost independent. On the contrary if $\rho_\tau \simeq d\tau^{-\beta}$, $0 < \beta < 1$, as $\tau \to 0$ then $\tau\rho_\tau \to 0$ and it can be easily shown that the process exhibits long range dependence. Indeed, we can show that

$$\sum_{|j|\leq n} E\{\varepsilon_k\varepsilon_{k+j}\} \simeq r(0) + \frac{2}{\tau\rho_\tau}\int_0^{n\tau\rho_\tau} r(t)dt \,, \tag{5}$$

provided that $\int_0^\infty |r(t)|dt$ exists. Thus the sum of the covariances of $\{\varepsilon_k\}$ may be unbounded if $\tau\rho_\tau \to 0$, i.e., we obtain the LRD noise process. It is also worth noting that if $\rho_\tau \simeq d\tau^{-1}$ then we have $E\{\varepsilon_k\varepsilon_{k+j}\} = r(dj)$ and thus we obtain the first noise model described in Assumption 1. On the contrary if $\rho_\tau = d$ then $E\{\varepsilon_k\varepsilon_{k+j}\} = r(dj\tau)$ and no consistency can be expected in this case.

In this paper we assess the accuracy of the signal reconstruction method for the both aforementioned noise models. Note that in [4] algorithms for signal recovery from samples observed in the presence of the linear SRD noise were only studied. Here we examine the statistical implications of the LRD assumption on the sampling and signal recovery problem. Then we apply our theory to the problem of designing consistent tests for verifying a hypothesis whether the signal f in model (2) belongs to some pre-specified finite dimensional subspace of the signal space. It is also worth mentioning that all presented results are obtained for the post-filtering reconstruction algorithm but they can be easily extended to other reconstruction methods examined in [2, 3, 4]. The prime goal of this work, however, is to evaluate the impact of various types of dependent errors on the choice of the sampling rate and on the accuracy of the reconstruction method. These results seem to be universal in the sense that other possible reconstruction algorithms will exhibit similar if not the identical behavior. In fact we show that in many special cases our rates agree with known optimal rates obtained in the signal processing and statistical literature.

2. RECONSTRUCTION ALGORITHMS FROM NOISY DATA

A naive reconstruction algorithm would use (1) with $\{f(k\tau)\}$ replaced by $\{y_k\}$ yielding $f_n(t) = \sum_{|k|\leq n} y_k \operatorname{sinc}(\pi\tau^{-1}(t-k\tau))$. It is easy to verify that in this case the global reconstruction error

$$\text{MISE}(f_n) = E\int_{-\infty}^{\infty}(f_n(t) - f(t))^2 dt \tag{6}$$

converges to infinity as $n \to \infty$ for any $\tau \leq \pi/\Omega$. This deficiency of $f_n(t)$ calls for a certain smooth correction of $f_n(t)$. This can be achieved, see [3, 4] for other alternatives, by filtering out in $f_n(t)$ all frequencies greater than Ω. Hence knowing only that $\Omega \leq W$ and applying an ideal low-pass filter with bandwidth W we obtain our basic reconstruction formula $\hat{f}_n(t) = f_n(t) * \sin(Wt)/\pi t$ which can be written in an explicit form as follows.

$$\hat{f}_n(t) = \tau \sum_{|k|\leq n} y_k \varphi(t - k\tau), \tag{7}$$

where $\varphi(t) = \sin(Wt)/\pi t$ is the reproducing kernel for $BL(\Omega)$.

In the next section we give conditions under which the $\text{MISE}(\hat{f}_n)$ converge to zero as $n \to \infty$ with a certain speed. It clear that the dependence influences only the stochastic part of the error, i.e., its integrated variance $IVAR(\hat{f}_n) = \int_{-\infty}^{\infty} E(\hat{f}_n(t) - E\hat{f}_n(t))^2 dt$. On the other hand, the bias term $\text{IBIAS}^2(\hat{f}_n) = \int_{-\infty}^{\infty}(Ef_n(t) - f(t))^2 dt$ can be evaluated in the similar way as in [4]. For the latter we require an assumption on the decay of $f(t)$ at $\pm\infty$, i.e., we need.

Assumption 3 *Let $f \in BL(\Omega)$ and let for $s \geq 0$ we have*

$$|f(t)| \leq c_1|t|^{-(s+1)}, \quad |t| > 0.$$

The issue of sampling representations analogous to (1) for non-band-limited signals is much more delicate. Generalized sampling theorems exist for some specific subspaces of $L_2(R)$, [1]. The results of this paper can be generalized to non-band limited signals by using the reproducing kernel $\varphi(t)$ with increasing W, such that $W = W_n \to \infty$ with a certain rate. In this case our algorithm requires not only optimal choice of the sampling period τ but also the bandwidth W. This issue is left for further studies.

3. ACCURACY ANALYSIS

Let us consider the global properties of $\hat{f}_n(t)$ for the both noise models introduced in Section 1. Let us start with the following universal expression for $IVAR(\hat{f}_n)$.

$$\mathrm{IVAR}(\hat{f}_n) = \frac{W}{\pi}\sigma_\varepsilon^2(2n+1)\tau^2 + \frac{2W}{\pi}(2n+1)\tau^2\sum_{j=1}^{2n}\left(1-\frac{j}{2n+1}\right) r(j)\,\mathrm{sinc}(j\tau W).$$

This explicit formula is essential for the sake of bounding $\mathrm{IVAR}(\hat{f}_n)$ for various dependent noise processes. At first we consider the noise model when the covariance function is independent of τ, i.e., the noise model characterized by Assumption 1. The following lemma gives the asymptotic behavior of the $\mathrm{IVAR}(\hat{f}_n)$ for the long memory noise processes of this kind.

Lemma 1 *Let Assumption 1 be met. Then for $\tau \to 0$ and $n\,\tau \to \infty$ we have*

$$\mathrm{IVAR}(\hat{f}_n) = (1+o(1))c(\alpha)(2n+1)W^\alpha\ \tau^{1+\alpha}\,L(1/\tau)$$

where $c(\alpha) = (\sin((\alpha+1)\pi/2)\Gamma(\alpha+1))^{-1}$.

Combining Lemma 1 with the results obtained in [3, 4] concerning $\mathrm{IBIAS}^2(\hat{f}_n)$ we can derive the following rate for $\mathrm{MISE}(\hat{f}_n)$.

Theorem 1 *Let a be positive constant and assume that the noise process obeys Assumption 1 with* $0 < \alpha < 1$.
Let f satisfy Assumption 3. Then selecting

$$\tau^* = an^{-\frac{2s+1}{2s+1+\alpha}}$$

we obtain

$$\mathrm{MISE}(\hat{f}_n) = O\left(n^{-\frac{2s\alpha}{2s+1+\alpha}}\right) .$$

It is worth noting that the sampling period is smaller for the LRD data than the SRD ones.
Let us now the noise model satisfying Assumption 2. As we have noticed in the discussion below Assumption 2 the case $\rho_\tau\,\tau \to 0$ as $\tau \to 0$ yields the LRD errors. Let us first examine the variance part of the reconstruction error.

Lemma 2 *Under Assumption 2 with $r(\bullet) \in L(0,\infty)$ and $\rho_\tau\,\tau \to 0$ as $\tau \to 0$ and $n\,\tau \to \infty$ we have*

$$\mathrm{IVAR}(\hat{f}_n) = \Big(1+o(1)\Big)\frac{2W}{\pi}(2n+1)\,\frac{\tau}{\rho_\tau}\int_0^\infty r(v)\,dv\,.$$

Combining the above lemma with the results obtained in [3, 4] we can derive the following rate for the τ dependent noise process.

Theorem 2 *Let a, d be positive constants and assume that the noise process obeys Assumption 2 with $r(\bullet) \in L(0,\infty)$ and $\rho_\tau = d\tau^{-\beta}$ with $0 < \beta < 1$.*
Let f satisfy Assumption 3. Then selecting

$$\tau^* = an^{-\frac{2s+1}{2s+1+\beta}}$$

we obtain

$$\mathrm{MISE}(\hat{f}_n) = O\left(n^{-\frac{2s\beta}{2s+1+\beta}}\right) .$$

It is worth noting that the rates in Theorem 2 can be arbitrary slow if β is approaching 0. In the extreme case when $\beta = 0$ there is no rate. This corresponds to the strong dependence of the correlation function on the sampling rate, i.e., when $E\{\varepsilon_k\varepsilon_{k+l}\} = r(l\tau)$ with $r(\bullet) \in L(0,\infty)$. It remains an open problem how to tackle this type of noise process. It is also expected that the rate obtained in Theorem 2 is slower than that in Theorem 1. In fact in order to compare these rates let us choose, e.g., $\alpha = 1/2$, $s = 1$. Some algebra shows that we have the rate $\mathrm{MISE}(\hat{f}_n) = O(n^{-2/7})$ in Theorem 1 and the rate $\mathrm{MISE}(\hat{f}_n) = O(n^{-1/4})$ in Theorem 2. Clearly the latter rate is slower than the former one.

4. LACK-OF-FIT TESTS

The results obtained in the previous section form the basis for designing a non-parametric lack-of-fit test. Hence we wish to test the null hypothesis $H_0:\ f = f_0$, where f_0 is a fixed band-limited signal satisfying Assumption 3, using a statistic based on the estimate $\hat{f}_n(t)$. We shall examine a test statistic which is the L_2 distance between $\hat{f}_n$ and f_0, i.e.,

$$D_n = \int_{-\infty}^{\infty}\left(\hat{f}_n(t) - f_0(t)\right)^2 dt.$$

To design the formal test based on D_n we need to find the limiting distribution of the test statistic. The problem of constructing consistent non-parametric lack-of-fit tests for the functional form of a signal has ben rarely addressed in the signal processsing literature and the first attempt was made in [7] where the noise process $\{\varepsilon_k\}$ is assumed to be of the iid type.

We extend these results to the case of correlated errors. Nevertheless, we consider the linear noise process $\varepsilon_k =$

$\sum_{j=0}^{\infty} \lambda_j Z_{k-j}$, where $\{Z_j\}$ is a sequence of i.i.d. random variables with $EZ_j = 0$, $\operatorname{var} Z_j < \infty$ and $E(Z_1^4) < \infty$. This is due to the fact that our technical developments relay heavily on the present known theory of central limit theorems for quadratic forms of dependent processes [8] . Let $\phi(\omega) = \sigma_Z^2 (2\pi)^{-1} \left|\hat{\lambda}(\omega)\right|^2$ be the power spectral density of $\{\varepsilon_k\}$, where $\hat{\lambda}(\omega)$ be the Fourier transform of $\{\lambda_k\}$. The main assumption in our derivation of the central limit theorem is the asymptotic behavior of $\phi(\omega)$ at $\omega = 0$. Thus we require that $\phi(\omega)$ is continuous at 0 or that for some $0 < \alpha < 1/2$

$$\phi(\omega) \simeq \omega^{-\alpha} L^*(1/\omega) \quad \text{as } \omega \to 0+\,, \tag{8}$$

where $L^*(\bullet)$ is a slowly varying function at infinity. Then it can be shown that under condition (8) and when the null hypothesis is true we have

$$(D_n - EI_n)/v_n \xrightarrow{d} N(0,1)\,. \tag{9}$$

where $N(0,1)$ is the standard normal random variable, $EI_n = \mathrm{IVAR}(\hat{f}_n)$ and

$$v_n^2 = \begin{cases} 16\,\pi\,\Omega\,n\,\tau^3\,\phi^2(0)(1+o(1)) \\ \quad \text{if } \phi(\omega) \text{ is continuous at } 0, \\ \frac{16\,\pi\Omega^{1-2\alpha}}{1-2\alpha}\, n\tau^{3-2\alpha} L^{*2}(1/(\Omega\tau))(1+o(1)) \\ \quad \text{if } \phi(\omega) \text{ meets (8) with } 0<\alpha<1/2. \end{cases} \tag{10}$$

On the other hand if the alternative hypothesis holds, i.e., that f_0 is not a true signal then we have that $(D_n - EI_n)/v_n$ tends to infinity.

All the aforementioned considerations yield the practical way of testing the hypothesis $H_0 : \; f = f_0$. Indeed for the selected confidence level $0 < \delta < 1$ we reject H_0 if

$$\frac{D_n - EI_n}{v_n} > F_N^{-1}(1-\delta)\,, \tag{11}$$

where $F_N^{-1}(1-\delta)$ is the upper $1-\delta$ quantile of the $F_N(\bullet)$. If the inequality in (11) is not valid we accept the hypothesis H_0. There are a number of computational issues related to the proposed test and this will be examined elsewhere.

5. CONCLUDING REMARKS

In this paper a thorough analysis of the post-filtering signal reconstruction method calculated from sampled data observed in the presence long-memory errors was given. The obtained result reveals that the rate of convergence can be arbitrary slow. To alleviate this problem one can apply higher oversampling rate. Yet another promising alternative would be to use random sampling, i.e., replace the sampling points $\{j\tau, |j| \le n\}$ by random points $\{\pi(j)\tau, |j| \le n\}$, where π is a randomly chosen permutation from a class of all permuations of the set $\{-n, \ldots, n\}$. Then the estimate $\hat{f}_n(t)$ would take the following form $\tilde{f}_n(t) = \sum_{|k|\le n} y_k \varphi(t - \pi(k)\tau)$, where y_k's are the observations taken at the nonrandom point $\{k\tau\}$. We conjecture that recovering methods based on random sampling can have an improved rate of convergence in the case of long-range dependent errors. In fact the rate $O(n^{-\frac{2s\alpha}{2s+1+\alpha}})$ obtained in Theorem 1 is expected to be replaced by a faster rate $O(\max(n^{-\alpha}, n^{-\frac{s}{s+1}}))$. The problem of establishing exact rates of signal recovery methods utilizing random sampling is left for future research. We also refer to [9] for some discussion of this issue in the context of the classical non-parametric regression estimation problem.

6. REFERENCES

[1] M. Unsner, “Sampling – 50 years after Shannon,” *Proceedings of the IEEE*, vol. 88, pp. 569–587, April 2000.

[2] M. Pawlak and E. Rafajlowicz, “On restoring band-limited signals,” *IEEE Trans. Information Theory*, vol. 40, pp. 1490–1503, 1994.

[3] M. Pawlak and U. Stadtmüller, “Statistical aspects of sampling for noisy and grouped data,” in *Advances in Shannon Sampling Theory: Mathematics and Applications*, J. Benedetto and P. Ferreira, Eds., pp. 317–342. Birkhauser, Boston, 2001.

[4] M. Pawlak, E. Rafajłowicz, and A. Krzyżak, “Post-filtering versus pre-filtering for signal recovery from noisy samples,” *IEEE Trans. Information Theory*, vol. 49, December 2003.

[5] P. Doukhan, G. Oppenheim, and M. Taqque, Eds., *Theory and Applications of Long-Range Dependence*, Birkhauser, 2003.

[6] O. Cappe, E.Moulines, J.C. Pesquet, A.Petropulu, and X.Yang, “Long-range dependence and heavy-tail modeling for teletrafic data,” *IEEE Signal Processing Magazine*, vol. 19, pp. 14–27, 2002.

[7] N.Bissantz, H.Holzmann, and A.Munk., “Testing parametric assumptions on band- or time-limited signals under noise,” *IEEE Trans. Information Theory*, vol. 51, pp. to appear, 2005.

[8] L.Giraitis and D.Surgailis, “A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotical normality of whittle’s estimate,” *Probability Theory and Related Fields*, vol. 86, pp. 87–104, 1990.

[9] S. Efromovich, “How to overcome the curse of long-memory errors,” *IEEE Trans. Information Theory*, vol. 45, pp. 1735–1741, July 1999.