SIGNAL SAMPLING AND RECOVERY WITH LONG-RANGE DEPENDENT ERRORS

M. Pawlak

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

ABSTRACT

We consider the extension of the Whittaker-Shannon interpolation reconstruction formula to the case of band-limited signals observed in the presence of correlated noise. Observing that in this situation the classical sampling expansion gives inconsistent reconstruction we apply the post-filtering strategy yielding a smooth correction of the interpolation series. We assess the accuracy of the method by the global L_2 error. A large class of dependent noise processes is taken into account. This includes short and long memory errors. Whereas the short memory errors have relatively small influence on the reconstruction accuracy, the long-memory errors can dramatically slow down the convergence rate. We explain this phenomenom by evaluating the speed at which the reconstruction error tends to zero.

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], [2]. 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 reconstructed from its discrete values $\{f(k\tau)\}$ by

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

provided that $\tau \le \pi/\Omega$, where $\mathrm{sinc}(t) = \sin(t)/t$. 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. Relatively little attention, however, has been given to statistical aspects of the sampling theorem, i.e., to the statistical analysis of (1) when only a finite record of noisy data is available. This important issue has been mentioned often in the signal processing literature but no algorithms with established convergence properties for a signal reconstruction from sampled and noisy data were given. The first rigorous theoretical treatment of this problem has been given in [3] and next in [4], see also [5] for an overview of various modifications of (1) to the case of noisy data. In all these contributions the white noise case has been mostly examined. In [6], however, the extension of the previous theory to short memory noise processes was also obtained. 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 [7]. Furthermore, an aggregation of short-memory processes can lead to long-range dependency effects.

In this paper we consider the following statistical model. One observes $N = 2n+1$ data points

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

and wishes to design a reconstruction scheme resembling (1) such that a certain reduction of the noise present in the data in obtained. Here $\{\varepsilon_k\}$ is the zero mean finite variance noise process. The following assumption on $\{\varepsilon_k\}$ is employed throughout the paper.

Assumption 1 *The noise $\{\varepsilon_k\}$ is a weakly stationary stochastic process with $E\varepsilon_k = 0$, var$(\varepsilon_t) = \sigma^2$, and cov$(\varepsilon_k, \varepsilon_\ell) = r(|k-\ell|)$, such that $\sigma^2 < \infty$ and that for $|k| \ge \delta > 0$ we have*

$$r(|k|) = c_r|k|^{-\alpha}, \quad 0 < \alpha \le 1. \qquad (3)$$

The above assumption implies that the power spectral density $h(\omega)$ of $\{\varepsilon_k\}$ has a pole at the origin, i.e., that $h(\omega) \approx c_h|\omega|^{-(1-\alpha)}$ as $\omega \to 0$. 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). An important example of dependent noise is the linear process

$$\varepsilon_k = \sum_{j=0}^{\infty} \lambda_j Z_{k-j}, \qquad (4)$$

where $\{Z_j\}$ is a sequence of iid random variables with zero mean and finite variance. Let $\lambda_j = c_\lambda j^{-p}$, $j = 1, 2, \ldots$, with $p > 0$. If $p > 1$ the process in (4) is SRD, whereas if $1/2 < p < 1$ then (4) is a stationary process with LRD. In fact, it can be shown that the covariance $r(k) = \sum_{\ell=0}^{\infty} \lambda_\ell \lambda_{\ell+k}$ is of order $c_r k^{-(2p-1)}$. We refer to [7] for an extensive overview of the theory and applications of long-range dependent processes. Nevertheless, to the best knowledge of the author there have been no attempts to study the sampling problem in the presence of the LRD noise.

In [6] algorithms for signal recovery from samples observed the presence of the linear SRD noise were studied. In this paper we examine the statistical implications of the LRD assumption on the sampling problem.

2. RECONSTRUCTION ALGORITHMS FROM NOISY DATA

A naive reconstruction algorithm would use (1) with $\{f(k\tau)\}$ replaced by $\{y_k\}$ yielding $f_n(t) = \sum_{|k|\le n} y_k \mathrm{sinc}(\pi\tau^{-1}(t-k\tau))$. It is easy to verify that the L_2 reconstruction error

$$MISE(f_n) = E\int_{-\infty}^{\infty} (f_n(t) - f(t))^2 dt \qquad (5)$$

tends to infinity as $n \to \infty$ for any $\tau \le \pi/\Omega$. This deficiency of $f_n(t)$ calls for a certain smooth correction of $f_n(t)$. This can be achieved, see [5], [6] for other alternatives, by filtering out in $f_n(t)$ all frequencies greater than Ω. Hence knowing only that $\Omega \le W$ 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.

$$\hat{f}_n(t) = \tau \sum_{|k|\le n} y_k \varphi(t - k\tau), \qquad (6)$$

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

In the next section we give conditions under which the $MISE(\hat{f}_n)$ converges to zero as $n \to \infty$ with a certain speed. We observe that the rate for the SRD case is not altered by the presence of correlation in the data. This is not the case for the LRD noise when we may observe a dramatic reduction of the rate. It is clear that the dependence influences only the stochastic part of the error, i.e. $IVAR(\hat{f}_n) = E\int_{-\infty}^{\infty}(f_n - Ef_n(t))^2 dt$. The bias term $IBIAS(\hat{f}_n) = \int_{-\infty}^{\infty}(Ef_n(t) - f(t))^2 dt$ can be evaluated in the similar way as in [4], [6]. For the latter we require an assumption on the decay of $f(t)$ at $\pm\infty$, i.e., we need.

Assumption 2 *Let $f \in BL(\Omega)$ and let for $s \ge 0$ we have*

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

3. ACCURACY ANALYSIS

Our first result gives a bound for $IVAR(\hat{f}_n)$ under Assumption 1. For the comparison let us recall, see [6], that $IVAR(\hat{f}_n)$ for the SRD noise is given by

$$IVAR(\hat{f}_n) \le \frac{W}{\pi}\{\sigma^2 + 2\sum_{\ell=1}^{\infty} |r(\ell)|\} N\tau^2. \qquad (7)$$

Note that for the white noise case we have the equality in (7) with the right-hand side of (7) replaced by $\frac{W}{\pi}\sigma^2 N\tau^2$.

Theorem 1 *Let $\{\varepsilon_k\}$ satisfy Assumption 1. Then for any $n \ge 1$ we have*

$$IVAR(\hat{f}_n) \le \frac{W}{\pi}\sigma^2 N\tau^2 + \frac{2(\alpha+1)}{\pi\alpha} \{W\sigma^2 c_r^{1/\alpha}\}^{\alpha/(\alpha+1)} N\tau^{\frac{2\alpha+1}{\alpha+1}}. \qquad (8)$$

It should be noted that the second term in (8) describes the additional error due to the presence of the LRD noise of order α. Furthermore if $\tau = c_\tau N^{-\kappa}$ for $\kappa > (\alpha + 1)/(2\alpha+1)$ then $\lim_{n\to\infty} IVAR(\hat{f}_n) = 0$. To get some insight into the behavior of (8) let us consider an important example of the LRD process $\{\varepsilon_k\}$ being the samples of the fractional Gaussian noise [8]. The bound in (8) was evaluated in this case with $W = \pi$. Figure 1 depicts the bound for $IVAR(\hat{f}_n)/\sigma^2 N$ versus the memory parameter $0 < \alpha \le 1$, for several values of sampling rate. It is apparent that oversampling can reduce the effect of the noise memory.

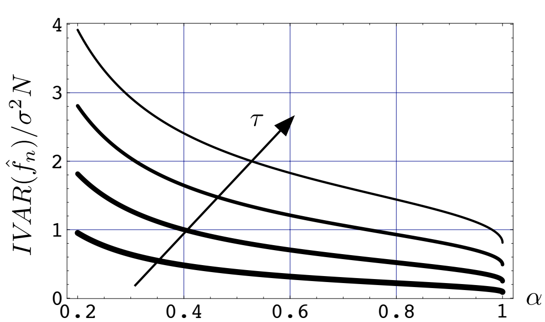

Fig. 1. $IVAR(\hat{f}_n)/\sigma^2 N$ versus α for the fractional Gaussian noise, $\tau = 0.3, 0.5, 0.7, 0.9$.

The result obtained in [6] suggest that under Assumption 2 $IBIAS(\hat{f}_n) = O((N\tau)^{-2s})$. This and Theorem 1 yield the following result concerning the rate of convergence.

Theorem 2 *Let Assumptions 1 and 2 be met. Then for the choice*

$$\tau^* = aN^{-\frac{(2s+1)(\alpha+1)}{2\alpha+1+2s(\alpha+1)}}$$

we have

$$MISE(\hat{f}_n) = bN^{-\frac{2s\alpha}{2\alpha+1+2s(\alpha+1)}}. \qquad (9)$$

The rate of convergence for the SRD noise is $O(N^{-\frac{s}{s+1}})$ with the sampling rate selected as $\tau^*_{SRD} = aN^{-\frac{2s+1}{2(s+1)}}$. This is clearly a faster rate than the one given in (9). It is also worth noting that τ^*_{SRD} is larger than τ^* specified in Theorem 2. The rate in (9) improves with s meaning that signals which more concentrate in the time domain are easier to estimate than those which have heavy tails. Figure 2 illustrates this point by plotting the exponent $2s\alpha/\{2\alpha+1+2s(\alpha+1)\}$ in (9) versus the memory parameter α for $s = 1, 3, 10$.

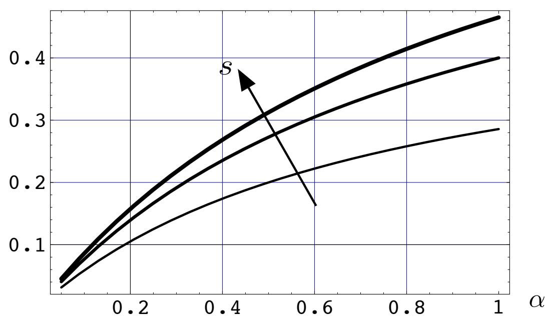

Fig. 2. The exponent in (9) versus α for $s = 1, 3, 10$.

Remark 1 *Theorem 2 specifies the optimal sampling rate τ^*. This choice depends on the noise memory parameter α and the signal tail parameter s. The problem of estimation α has been extensively examined in the time series literature [9], [7]. In many cases one can also recover the tail parameter. This could lead to an adaptive choice of the sampling rate.*

Remark 2 *The results of this paper can be generalized to the case when the correlation function depends on τ, i.e., $cov(\varepsilon_k, \varepsilon_{k+\ell}) = r_\tau(|\ell|)$ and $r_\tau(|\ell|) = r(|\ell|)\rho(\tau)$ for $r(|\ell|)$ satisfying Assumption 1 and $\lim_{\tau\to 0}\rho(\tau) = c$, some constant c. If, however, $r_\tau(|k|) = \rho(\tau|k|)$ with $\rho(t) = c_\rho|t|^{-\alpha}$, $\alpha > 0$, then we obtain that $IVAR(\hat{f}_n) = O(N\tau)$ and consequently no convergence of the error to zero is possible.*

4. CONCLUDING REMARKS

In this paper a thorough analysis of the post-filtering signal reconstruction method calculated from sampled data observed in the presence of the long-memory noise was given. The obtained result, see (9), 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 (2) by $y_k = f(\tau_k) + \varepsilon_k$, where $\{\tau_k\}$ is a sequence of N random time points over a certain interval centered at the origin. Then the estimate in (6) would take the following form

$$\hat{f}_n(t) = \sum_{|k|\le n} (\tau_{(k)} - \tau_{(k-1)}) y_{[k]} \varphi(t - \tau_{(k)}),$$

where $\tau_{(-n-1)} < \tau_{(-n)} < \cdots < \tau_{(n)}$ is the ordered version of $\{\tau_k\}$ and $y_{[k]}$'s are the observations paired with $\tau_{(k)}$'s. The reason that this estimate can have an improved performance is due to the fact that the dependence between randomly mixed data is smaller than between consecutive observations as this is the case in the deterministic sampling.

5. REFERENCES

[1] J. R. Higgins, *Sampling Theory in Fourier and Signal Analysis*, Clarendon Press, Oxford, 1996.

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

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

[4] M. Pawlak and U. Stadtmüller, “Recovering band-limited signals under noise,” *IEEE Trans. Information Theory*, vol. 42, pp. 1425–1438, 1996.

[5] 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.

[6] 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.

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

[8] B. B. Mandelbrot and J. W. Van Ness, “Fractional Brownian motions, fractional noises and applications,” *SIAM Rev.*, vol. 10, pp. 422–437, 1968.

[9] J. Gewke and S. Porter-Hudak, “The estimation and applications of long-memory time series models,” *Journal of Time Series*, vol. 4, pp. 221–237, 1983.