Iranian Journal of Information Science and Technology

SPECTRAL ANALYSIS OF SEQUENTIALLY SAMPLED SIGNALS FOR APPLICATIONS IN INFORMATION PROCESSING

N. FAGHIH, Ph.D.
Department of Management and Accounting
College of Social Sciences, Shiraz University
Shiraz, I. R. of Iran

Abstract - This paper studies the sequential sampling scheme, as a solution to the problem of aliasing, where the sampling interval is restricted to a minimum allowable value ΔT. In the sequential sampling, the signal is sampled at intervals of ΔT, $\Delta T+\Delta\tau$, $\Delta T+2\Delta\tau$, $\Delta T+3\Delta\tau$, ...; where $\Delta\tau < \Delta T$ and may be selected as desirable. The sequential sampling is, however, analyzed and it is proven that, when the ratio $\Delta T/\Delta\tau$ is an integral number, the associated spectral estimates give a Nyquist frequency $\frac{1}{2\Delta\tau}$. This sampling scheme can, therefore, be employed to yield a required cut-off frequency. The autocorrelation function estimation from the sequentially sampled data is then considered and the approach to this is discussed. Simulation studies are also used for empirical investigations of the sequential sampling. Furthermore, since with this sampling scheme no estimates of the autocorrelation coefficients are available in the initial time interval from 0 to ΔT, the initial autocorrelation coefficients estimation method is applied to solve this problem. This, in addition, provides some empirical studies of the latter approach. Moreover, the application of the autocorrelation function extrapolation method is considered, as a means of minimizing the sampling time and costs. The contribution of the sequential sampling in obtaining a desired cut-off frequency, is also demonstrated by data simulation.

Keywords - Sequential Sampling, Spectral Analysis, Autocorrelation Function, Aliasing, Fourier Transform.

INTRODUCTION

Some data acquisition systems have a minimum allowable sampling interval and do not provide a desired sampling period less than a minimum allowable value. This may be due to some restrictions set by the measuring instrument that has to be used. [20, 19, 2, 3, 1, 8, 5]

Let the minimum allowable sampling time be ΔT; if the uniform sampling scheme is employed, then the Nyquist or cut-off frequency is known [4] to be given as:

$$f_c = \frac{1}{2\Delta T} \qquad (1)$$

This would mean that if frequencies higher than f_c are present, aliasing will occur. Otherwise, the signal would have to be filtered so that only frequencies below f_c are passed and, therefore, the spectral analysis will be restricted [4].

The sequential sampling scheme can, however, be employed to obtain an autocorrelation function with estimates $\Delta\tau$ apart, where $\Delta\tau < \Delta T$, with the exception of

coefficients lying inside the range $R(0) \rightarrow R(\Delta T)$. In this sampling scheme, the signal would be sampled at intervals of:

$\Delta T, \Delta T+\Delta\tau, \Delta T+2\Delta\tau, \Delta T+3\Delta\tau, \ldots$

From the sampled signal, an autocorrelation function can be obtained with the coefficients:

$R(0), R(\Delta T), R(\Delta T+\Delta\tau), R(\Delta T+2\Delta\tau), \ldots$

While ΔT is restricted, the value of $\Delta\tau$ may be chosen as desirable. In this paper, it will be proven that the sequential sampling can give an increased cut-off frequency as:

$$f_{cs} = \frac{1}{2\Delta\tau} \qquad (2)$$

The sequential sampling can, therefore, be employed to overcome aliasing and the restrictions of spectral analysis, by selecting a sufficiently small value for $\Delta\tau$.

The missing coefficients, inside the range $R(0) \rightarrow R(\Delta T)$ may also be given estimates by using the initial autocorrelation coefficients estimation method. A discrete autocorrelation function, with an equi-spaced interval $\Delta\tau$, is then provided; this can give spectral estimates, subject to the discrete Fourier cosine transformation.

The first and second order Gaussian processes with prescribed autocorrelation functions were used for simulation studies of the sequential sampling and will be reported in this paper. As already stated, the initial coefficients estimation method is also applied to estimate the missing coefficients. This would, in turn, provide an empirical investigation of the latter method.

The incorporation of the autocorrelation function extrapolation approach with the sequential sampling scheme, is also considered. As will be seen, the sequential sampling requires a lengthy sampling time and the extrapolation method can assist in reducing this. A portion of the autocorrelation function is, therefore, estimated from the data and used to obtain the missing lag values; the function is then extrapolated to decay. The discrete Fourier cosine transformation is applied to estimate the spectrum from this function.

The contribution of the sequential sampling in minimizing the problem of aliasing is also demonstrated by a simulation example.

THE CUT OFF FREQUENCY IN THE SEQUENTIAL SAMPLING

The cut-off frequency provided by the sequential sampling scheme is considered in this section. The analysis employs the impulse representation of a continuous signal as an approach to discretization [15, 13, 11, 7, 10, 16].

In the sequential sampling, the signal is sampled at intervals of:

$\Delta T, \Delta T+\Delta\tau, \Delta T+2\Delta\tau, \Delta T+3\Delta\tau, \ldots$

The sampling instants are, therefore, given by:

$$t_i = 0, \quad \Delta T, 2\Delta T+\Delta\tau, 3\Delta T+3\Delta\tau, 4\Delta T+6\Delta\tau, \ldots \qquad (3)$$

This can be written as:

$$t_i = i\Delta T + [\sum_{r=0}^{i} r - i]\Delta\tau, \quad i = 0,1,2,3,\ldots \qquad (4)$$

Since, $\sum_{r=0}^{i} r = \frac{i}{2}(i+1)$, then equation (4) gives:

$$t_i = i\Delta T + \frac{i}{2}(i-1)\Delta\tau \qquad (5)$$

When a continuous signal x(t) is sampled, the sample values $x(t_i)$ are acquired. A discrete autocorrelation function, with coefficients $R(\tau_j)$, may be obtained from the discrete signal, by contributions of the products $x(t_i)x(t_{i+j})$. Equation (5) can be used to give the time delay τ_j as:

$$\tau_j = t_{i+j} - t_i = \Delta T + [\frac{j}{2}(j-1) + ij + (j-1)\mu]\Delta\tau \qquad (6)$$

$$i = o, 1, 2, \ldots$$

$$j = o, 1, 2, \ldots$$

where μ is a constant given by:

$$\mu = \Delta T/\Delta\tau \qquad (7)$$

It is seen from equation (6) that for j=0, the time delay is zero and for j=1, the time delay is $\Delta T+i\Delta\tau$ (where i=0, 1, 2, 3, ...). An autocorrelation function is, therefore, obtainable at discrete values of the time delay given as:

$$\tau_n = o, \Delta T + n\Delta\tau \quad n = 0, 1, 2, 3, \ldots \qquad (8)$$

If the ratio μ is an integral number, then higher values of j would also provide more contributions to the autocorrelation estimates at the above time delays τ_n. This is because j(j-1)/2 is always even, and any value of j would hence add a multiple of $\Delta\tau$ to ΔT.

The discrete autocorrelation function may be represented as:

$$R^*(\tau) = \Delta\tau . R(\tau)\delta_b(\tau) \qquad (9)$$

where $R(\tau)$ is the continuous autocorrelation function and $\delta_b(\tau)$ is the following form of the delta comb:

$$\delta_b(\tau) = \delta(\tau) + \delta(\tau - \Delta T) + \delta(\tau - \Delta T - \Delta\tau) + \delta(\tau - \Delta T - 2\Delta\tau) + L \qquad (10)$$

It is established [15, 13] that the Fourier transform of equation (10) can be written as:

$$\Delta_b(\omega) = 1 + e^{-j\omega\Delta T}\sum_{0}^{\infty} e^{-j\omega n\Delta\tau} \qquad (11)$$

which by manipulation [15, 13] can be re-written as:

$$\Delta_b(\omega) = \frac{1 - e^{-j\omega\Delta\tau} + e^{-j\omega\mu\Delta\tau}}{1 - e^{-j\omega\Delta\tau}} \tag{12}$$

where substitution has also been made for ΔT from equation (7).

The Fourier transformation of $R^*(\tau)$ gives the spectral density $S^*(\omega)$ corresponding to the sampled signal and that of $R(\tau)$ would yield the spectral density $S(\omega)$ of the original continuous signal. The approach adopted for the Fourier transformation of equation (9) is based on the convolution and residue theorems [15]. By evaluating the residue terms [15] and using the convolution property [15], for substitution into equation (9), the Fourier transform of this equation ca be obtained as:

$$S^*(\omega) = \sum_{-\infty}^{\infty} e^{j2\pi n\mu} S(\omega + 2n\omega_{cs}), \quad \omega_{cs} = \pi_{/\Delta\tau} \tag{13}$$

However, if the ratio $\mu=(\Delta T/\Delta\tau)$ is an integral number then,

$$e^{j2\pi n\mu} = 1$$

noting that n is also an integer. Substituting this into equation (13) gives:

$$S^*(\omega) = \sum_{-\infty}^{\infty} S(\omega + 2n\omega_{cs}), \quad \omega_{cs} = \pi_{/\Delta\tau} \tag{14}$$

Now, consider the periodicity of $S^*(\omega)$; this can also be examined by applying the corresponding methods [15, 13, 11, 7, 10, 16]. Using equations (9), (10), (11) and the rules established for discrete Fourier transformation [15, 13], it can be written:

$$\Delta\tau[R(0) + e^{-j\omega\Delta T}\sum_{0}^{\infty} R(\Delta T + n\Delta\tau)e^{-j\omega n\Delta\tau}] = \Delta\tau[R(0) + e^{-j\omega\mu\Delta\tau}\sum_{0}^{\infty} R(\Delta T + n\Delta\tau)e^{-j\omega n\Delta\tau}] \tag{15}$$

and then for an integer m:

$$\Delta\tau[R(0) + e^{-j(\omega+2m\omega_{cs})\mu\Delta\tau}\sum_{0}^{\infty} R(\Delta T + n\Delta\tau)e^{-j(\omega+2m\omega_{cs})n\Delta\tau}] =$$

$$\Delta\tau[R(0) + e^{-j\omega\mu\Delta\tau}e^{-j2m\omega_{cs}\mu\Delta\tau}\sum_{0}^{\infty} R(\Delta T + n\Delta\tau)e^{-j\omega n\Delta\tau}.e^{-j2m\omega_{cs}n\Delta\tau}] =$$

$$\Delta\tau[R(0) + e^{-j\omega\mu\Delta\tau}.e^{-j2m\mu\pi}\sum_{0}^{\infty} R(\Delta T + n\Delta\tau)e^{-j\omega n\Delta\tau}.e^{-j2mn\pi}] =$$

$$\Delta\tau[R(0) + e^{-j\omega\mu\Delta\tau}.e^{-j2m\mu\pi}\sum_{0}^{\infty} R(\Delta T + n\Delta\tau)e^{-j\omega n\Delta\tau}] \tag{16}$$

since m and n are integers. If μ is also an integral number, this would reduce to:

$$\Delta\tau[R(0) + e^{-j\omega\mu\Delta\tau}\sum_{0}^{\infty} R(\Delta T + n\Delta\tau)e^{-j\omega n\Delta\tau}]$$

from which it follows that:

$$S^*(\omega + 2m\omega_{cs}) = S^*(\omega) \tag{17}$$

This is the mathematical statement for $S^*(\omega)$ to be periodic with period $2\omega_{cs}$. Otherwise, if μ is not an integral number,

$$S^*(\omega + 2m\omega_{cs}) \neq S^*(\omega) \tag{18}$$

and the requirement for periodicity is not met.

It is, therefore, seen that when the ratio $\mu(=\Delta T/\Delta\tau)$ is an integral number, the periodic pattern, conforming with the Nyquist theorem, [15, 16] is obtained. That is, the sequential sampling gives a cut-off frequency $\omega_{cs} = \pi_{/\Delta\tau}$ or $f_{cs} = \frac{1}{2\Delta\tau}$. On the contrary, when μ is not a whole number, $S^*(\omega)$ is related to the true spectral density by equation (13); it includes a complex term and is not periodic.

THE AUTOCORRELATION ESTIMATION FROM SEQUENTIALLY SAMPLED SIGNALS

Consider the zero mean ergodic data being sampled sequentially. It was already noted that a discrete autocorrelation function, with the coefficients:

$$R(0), R(\Delta T), R(\Delta T + \Delta\tau), R(\Delta T + 2\Delta\tau), \ldots$$

is obtainable from these data. An autocovariance function with discrete values at time delays (t_m - t_n), can be estimated from the contributions of the products $x(t_n)x(t_m)$ of the sample values; the autocovariance function may be normalized to yield the autocorrelation coefficients [4]. At least one contribution per estimate would be available but if the ratio $\mu(=\Delta T/\Delta\tau)$ is an integral number, then more contributions may also be available for some of the estimates. Improvement may, however, be gained upon the estimates by repeating the sample sequences and averaging over the contributions from different sequences.

The number of sample values required in each sequence, is determined by the maximum desirable time delay. Let this maximum time delay be $M\Delta\tau$, that is, the last autocorrelation coefficient be $R(M\Delta\tau)$. This time delay can be written as:

$$M\Delta\tau = \Delta T + (M - \mu)\Delta\tau \qquad (19)$$

since, $\Delta T=\mu\Delta\tau$. In the sequential sampling, after the sample value x(0), each sample value x(i) delays relative to its precedent by $\Delta T+(i-1)\Delta\tau$. This indicates that, including x(0), the number of sample values required in each group would be:

$$N_g = M-\mu+2 \qquad (20)$$

In other words, as given by equation (5), a maximum sampling instant:

$$t_m = (M - \mu + 1)\Delta T + \frac{1}{2}(M - \mu + 1)(M - \mu)\Delta\tau \qquad (21)$$

would be required. The sampling instants above t_m would give time delays larger than the maximum required value ($M\Delta\tau$). As many contributions as desired for averaging are, however, acquirable only by repeating the sequence and using the time instants zero to t_m.

The estimation of the lag values can take place as follows. In each sequence, N_g data values $x_1, x_2, \ldots, x_{N_g}$ have to be sampled and stored. The corresponding sampling instants $t_1, t_2, \ldots, t_{N_g}$ may also be stored or generated, while required, from equation (5). The products $x_n x_{n+r}$, with $n=1, 2, \ldots N_r$ and $r=1, 2, \ldots L$, can be formed from the data; where N_r and L depend on M, i.e. the required maximum time delay. The procedure is analogous to that of the autocorrelation estimation from equi-spaced data. [4] The corresponding time delays (t_{n+r}-t_n) are also calculated from the stored values of

the sampling instants or from equation (6), before the products are formed. If this time delay does not exceed the maximum required value ($M\Delta\tau$), then it is divided by $\Delta\tau$ to find whether it yields an integral number, i.e. the lag number to which the product is contributing. It is noted that if μ is a whole number, then the division will always give an integral number. The product $x_n x_{n+r}$ is now formed and added to the previous state of the relevant array element of the corresponding lag number; the initial state can be set to zero when starting the computation. However, if the time delay exceeds $M\Delta\tau$, then the next value of n will be considered.

When a lag product is found contributing to the summation, the corresponding contribution counter is increased by unity. It is noted that the number of contributions to each lag value estimate, per sequence, is fixed and may be computed only once. It can be computed by solely considering the sampling instants and the number of sample values used in a sequence, as explained before.

The values of L and N_r would depend on the maximum time delay required and may be determined as follows. In the sequential sampling, the minimum time interval elapsing between the sampling instants would be ΔT, and, hence, the maximum useful value for L may not exceed $M\Delta\tau/\Delta T$. Replacing ΔT by $\mu\Delta\tau$ gives:

$$L = M/\mu \qquad (22)$$

If M/μ is not an integral number, then its integer part is taken for L. It is noted that any lag product $x_n x_{n+r}$ with r>L, would give a time delay exceeding the required maximum value ($M\Delta\tau$). Similarly, N_r can be given as:

$$N_r = N_g-(r-1)\,\mu-1 \qquad (23)$$

For any displacement r the product $x_n x_{n+r}$, with $n>N_r$, would correspond to a time delay exceeding $M\Delta\tau$.

The sums of the products $x_n x_{n+r}$ obtained in each sequence may be divided by the corresponding number of contributions in a sequence. These mean values from various sequences, are added together and divided by the number of sequences to give the final average values. Equivalently, the sums can be calculated over all the sequences, then divided by the total number of contributions corresponding to each lag number to give the mean values.

However, the resulting auto covariance function may, at the end, be normalized (dividing through by the value at zero time delay) to yield the autocorrelation coefficients. The fact that the normalization should be left to the end is considered and appreciated in the following section.

THE NORMALIZATION AND THE AVERAGING PROCESS

When data are acquirable in sequences of small size, the lag value estimates may be improved only by repeating the sequences and hence, increasing the total number of contributions. The question worth considering is whether the auto covariance function estimates from each sequence, should be averaged by repetition and then normalized at the end or, otherwise, the normalization should take place in each sequence and the

resulting autocorrelation function estimates averaged. It may be useful, not only in the sequential sampling but also in the uniform sampling, when a restricted storage capacity is available for computation and the data sampling has to take place in sequences.

This section considers the above empirically, by simulating the first and second order Gaussian processes, using uniform sampling.

Consider the first order Gaussian process with autocorrelation function:

$R(\tau) = \exp(-\tau)$

The data were simulated with a time interval $\Delta\tau=0.05$. In each sequence, 51 sample values were used to obtain contributions to the lag value estimates at time delays of zero to $50\Delta\tau$. The sequence was repeated 2 000 times and, hence, a total sample size of 102000 was employed. Firstly, normalization took place in each sequence and the resulting autocorrelations were used to obtain an average for each coefficient. The estimates are shown and compared with the true curve, in Figure 1. Secondly, the normalization did not take place in each sequence and was left to the end, namely after the average values were obtained from contributions of different sequences. The corresponding autocorrelation coefficient estimates are shown and compared with the true curve, in Figure 2. It is seen that the estimates in the second case are better than the first case. In fact, the estimates in the first case are not satisfactory at all. It is also observed that the accuracy decreases as the time delay is increased. This is understood by, for instance, considering R(50) and noting that there are only 2 000 contributions to the estimate of this coefficient, i.e. one per sequence; although the total sample size is 102 000.

Next, the second order Gaussian process with autocorrelation function:

$R(\tau) = \exp(-\tau)\cos\pi\tau$

was simulated and a sampling interval $\Delta\tau=0.05$ was used. The autocorrelation coefficients were estimated as in the previous example. Figure 3 shows the estimates corresponding to the case that the normalization took place in each sequence and Figure 4 displays those corresponding to normalizing only once at the end. The observations made from these figures are similar to those of the former example. They indicate that the normalization should take place once at the end and not in each sequence. This is also more efficient, since the normalization is not to be repeated for every sequence.

SIMULATION STUDIES OF THE SEQUENTIAL SAMPLING SCHEME

In this section, the first and second order Gaussian processes with known autocorrelation functions, by employing the methods of generation given in [17, 6], are used for simulation studies of the sequential sampling scheme. The data were generated (sampled) at the sequential times given by equation (3) and, were then used to estimate the autocorrelation coefficients according to the specifications already explained in this paper.

The missing coefficients lying inside the range $R(0) \rightarrow R(\Delta T)$ are also estimated by applying the initial autocorrelation coefficients estimation method (an account of which is presented in the appendix). The estimated discrete autocorrelation functions are

shown and compared with the true functions. Subsequently, the application of the autocorrelation function extrapolation method [18, 21] in the associated spectral estimation, is also considered. Finally, an example is provided to demonstrate the contribution of the sequential sampling scheme in reducing the problem of aliasing.

For the simulation studies reported in this paper, a ratio $\mu=5$ was assumed. This would give four missing lag values: R(1), R(2), R(3) and R(4). When a maximum coefficient R(50) is to be estimated, as seen from equation (20), 47 sample values would be required in each sequence. In each sequence then, there would be 14 lag values having only one contribution to each, 24 lag values having two contributions and, three contributions to each of 7 lag values. The coefficient R(45) only, would have five and R(0) would have 47 contributions per sequence. It is noted that although the total sample size (used by repeating the sequence) may be relatively large, but the total number of contributions to most of the lag value estimates would remain relatively small. For this reason, unfortunately, the sequential sampling scheme can become costly for accurate estimates. Here, the sequences have been repeated for different number of times; 15 000, 4 000 and 1 000 were used as repeating times.

Consider the first order Gaussian process with autocorrelation function:

$$R(\tau) = \exp(-\tau) \qquad (24)$$

The values of $\Delta T = 0.25$ and $\Delta\tau = 0.05$ were selected. The data were generated sequentially and the autocorrelation coefficients were estimated, by repeating the sequence 15 000 times. The coefficients $R(1) \rightarrow R(4)$, which could not be directly obtained from the data, were also given estimates by the initial coefficients estimation method, using the known values R(0) and $R(5) \rightarrow R(50)$. The resulting discrete autocorrelation function with a uniform time interval of $\Delta\tau = 0.05$, is shown and compared with the true curve in Figure 5. It is seen that the coefficients obtained directly from the data are plausible and lie either on the true curve or close to it. The initial coefficient estimates are also found satisfactory, but R(1) deviates slightly from the true value. This is a property of the interpolation method that estimation of R(1) is more susceptible to error, since it uses all the known and estimated initial values following it.

The repeating number was then reduced to N_s=4 000 and the discrete autocorrelation function was obtained. It is shown and compared with the true curve in Figure 6, where the sample size reduction has been reflected in the reduced accuracy of the values estimated directly from the data and also the initial coefficient estimates. However, the deviations of the estimates from the true values, do not seem to be very significant.

The repeating number was further reduced to N_s= 1 000 and the estimated discrete autocorrelation function is shown and compared with the true curve in Figure 7. It is seen that the accuracy of the estimates has been considerably reduced and, consequently, estimation of the missing values has not been possible. This is because, no values could be assigned to the initial coefficients to maintain the corresponding autocorrelation matrix non-negative definite, as a theoretical basis required by the interpolation method (outlined in the appendix). It is also seen that when this is the case, it would be notified by the estimating routine.

Another decaying exponential function was also considered. The second order Gaussian process with autocorrelation function:

$$R(\tau) = \frac{1}{2}[\exp(-\tau) + \exp(-2\tau)] \quad (25)$$

was simulated, using the values $\Delta T = 0.2$ and $\Delta\tau = 0.04$. Firstly, the repeating number N_s= 15 000 was used and the estimated function is shown and compared with the true curve in Figure 8. Secondly, the value N_s= 4 000 was used and Figure 9 shows the corresponding results. The observations made from these figures are the same as those in the previous (first order) case.

The second order Gaussian process with decaying cosinusoid autocorrelation function:

$$R(\tau) = \exp(-\tau)\cos\pi\tau \quad (26)$$

was also simulated. The values $\Delta T = 0.5$ and $\Delta\tau = 0.1$ were chosen for the sequential sampling. Again the same repeating numbers were used. Figure 10 shows the results corresponding to N_s=15 000, where the estimates obtained directly from the data are seen to be plausible. The estimated values of the missing coefficients R(4) and R(3) also lie on the true curve and those of R(2) and R(1) lie close to it, with the latter being less accurate than the former. This would be expected as already explained, noting that in addition, the imperfections of the values estimated from the data would also affect the estimation of the missing values.

Figure 11 also shows the results corresponding to N_s= 4 000. It is seen that some of the estimated values sustain small inaccuracies; these estimates have small number of contributions per sequence. The inaccuracies have also been reflected in the estimates of the missing initial coefficients; the estimates become less accurate from R(4) to R(1).

The effect of the above inaccuracies on the subsequent spectral estimation will be demonstrated in the following section.

THE AUTOCORRELATION FUNCTION EXTRAPOLATION METHOD AND THE SEQUENTIAL SAMPLING

The sequential sampling is a slow process, simply because the sampling intervals can not be less than ΔT and would increase arithmetically, as given by equation (5). The larger the maximum time delay required for the autocorrelation function to be estimated, the slower would be the sampling process. It is especially important to note that for a satisfactory estimation of the missing coefficients, accurate estimates have to be obtained from the data. This would require a large total sample size and would increase the sampling time and costs considerably.

However, in practice, for estimation of the missing coefficients, only a portion of the autocorrelation function can be sufficient. When estimates are given to the missing values, the resulting truncated function may then be extrapolated for spectral estimation. Therefore, this can provide a more accurate estimation of the autocorrelation function with less sampling time and costs.

This section considers the application of the autocorrelation function extrapolation

method [8, 21] to the sequential sampling. The same data with the same repeating numbers as already used in this paper, are simulated to estimate a portion of the autocorrelation function. The missing values are given estimates and the extrapolation method is then employed to extend the functions and estimate the spectra by cosine transformation [4] of the discrete extrapolated functions.

The first order process, with the autocorrelation function given by eqution (24), was simulated, using $\Delta T = 0.25$ and $\Delta\tau = 0.05$. The repeating number $N_s = 15\ 000$ was selected and the lag values R(0) and $R(5) \rightarrow R(16)$ were estimated; 13 sample values only were required in each sequence. The estimated lag values were employed to give estimates to the missing values $R(1) \rightarrow R(4)$. The truncated discrete function, $R(0) \rightarrow R(16)$, was used to extrapolate the function. The resulting discrete function, up to R(50), is shown and compared with the true curve in Figure 12. It is seen that the coefficients estimated from the data and the estimates given to the missing values, lie close to the true values. The extrapolated values deviate from the true values, but not very significantly. This is understood to be due to the inaccuracies of the estimates prior to extension. The spectrum was then estimated by cosine transformation of the discrete function extended up to R(200). This is shown and compared with the true spectrum in Figure 13. The estimated spectrum is observed to be in good agreement with the true spectrum and does not pronounce the small inaccuracies of the estimated autocorrelations.

The repeating number was then reduced to $N_s = 4\ 000$. Figure 14 shows the corresponding autocorrelation estimates and Figure 15 displays the resulting spectrum. It is seen that the autocorrelation estimates have become less accurate and the deviations of the extrapolated coefficients from their true values have increased. These inaccuracies have been reflected in the subsequent spectrum. The estimated spectrum is inaccurate and does not appear to decay smoothly, although no negative lobes are present.

The second order process, with the autocorrelation function given by equation (25), was also considered. The values $\Delta T = 0.2$ and $\Delta\tau = 0.04$ and the repeating numbers N_s=15 000 and N_s= 4 000 were used to estimate R(0), $R(5) \rightarrow R(16)$. Figures 16 and 17 show the results corresponding to N_s= 15 000 and figures 18 and 19 show the results corresponding to N_s= 4 000. These figures give the same observations as in the previous (first order) case.

The second order process with the decaying cosinusoid autocorrelation function, given by equation (26), was also simulated, using $\Delta T = 0.5$ and $\Delta\tau = 0.1$. The repeating number N_s= 15 000 was used and the lag values R(0) and $R(5) \rightarrow R(28)$ were estimated. In each sequence, 25 sample values would be required. The missing coefficients were obtained and extrapolation then took place from R(29). The resulting discrete autocorrelation function is shown and compared with the true curve, in Figure 20. It is seen to be satisfactory, but the estimates given to the missing values R(1) and R(2) are not very close to the true values. These two estimates reflect the small statistical inaccuracies of $R(5) \rightarrow R(28)$ and also, in addition, the small imperfections in the estimates given to the missing values R(3) and R(4).

The spectral estimates were obtained by extending the function up to R(200). Figure 21 compares the estimated spectrum with the true spectrum and shows that although it is not exact, it is a satisfactory representation of the latter, and the position and amplitude of the peak is fairly accurate. The estimated spectrum does not, however, decay smoothly beyond the peak and its value at zero frequency is slightly lowered. This may be attributed to the inaccuracies resulting in the estimates of the missing values R(1) and R(2).

Next the repeating number was, reduced to N_s = 4 000. The corresponding autocorrelation estimates are shown and compared with the true curve, as in Figure 22. It is observed that the estimates given to the missing initial coefficients are strongly pronouncing the statistical inaccuracies of the lag values obtained from the data, which would have increased as a result of reducing the sample size. The subsequently estimated spectrum is also shown and compared with the true spectrum in Figure 23, where it is seen to have become less accurate. Although the position and amplitude of the peak still appears to be satisfactory, the unevenness beyond the peak has increased and the value at zero frequency has lowered further.

Finally, to demonstrate the contribution of the sequential sampling in minimizing the aliasing problem, the Gaussian process with the following autocorrelation function was also simulated:

$$R(\tau) = \exp(-\tau)\cos 12\pi\tau \qquad (27)$$

The minimum allowable sampling time was assumed to be $\Delta T = 0.125$. When uniform sampling was used with this time interval, an aliased spectrum resulted. The discrete autocorrelation function estimated from the equi-spaced data is shown and compared with the true curve in Figure 24. The extrapolation method was then employed to obtaion the spectrum, which has a cut-off frequency of 4 (less than the true frequency content, namely 6) and is shown in Figure 25. The spectrum is seen to exhibit an aliased peak at f=2.

To minimize the aliasing problem, sequential sampling was then applied, using $\Delta\tau = 0.025$. The data were sampled sequentially and the autocorrelation coefficients R(0) and R(5) $\rightarrow$R(20) were estimated with a repeating number N_s= 15 000. The truncated discrete autocorrelation function has a uniform time interval $\Delta\tau = 0.025$, with the exception of the missing lag values R(1)$\rightarrow$R(4) (the ratio $\mu = 5$). Estimates were given to the missing values and the truncated function was extrapolated from R(21). The resulting discrete function up to R(50) is shown and compared with the true curve in Figure 26. The values estimated directly from the data are seen to be in good agreement with the true curve, and the estimates given for the missing values also conform with the latter. It is interesting to note that these values appear to maintain the true wavelength of the function. However, the extrapolated values do not appear to be so accurate. It seems that the slight statistical inaccuracies of the coefficients estimated from the data and imperfections of the estimates given to the missing values have been reflected in the extrapolated values.

The spectrum estimated by extending the discrete function up to R(200), is also shown and compared with the true curve, in Figure 27. It has a cut-off frequency of 20

and although not exact, it seems to be a satisfactory estimation for an experimentally obtained spectrum. Some negligibly small peaks are also observed at either side of the main peak. These small inaccuracies present in the spectrum may only be reduced by the willingness to expend a larger total record length (sampling time) and to estimate a larger portion of the autocorrelation function before extrapolating. It is, however, noted that a higher cut-off frequency has been obtained by applying the sequential sampling scheme, which would have not been possible in the uniform sampling, due to the minimum allowable sampling interval being restricted.

DISCUSSION AND CONCLUDING REMARKS

This paper has considered the sequential sampling scheme, as a solution to the problem of aliasing, where the sampling interval is restricted to a minimum allowable value ΔT. In the sequential sampling, the signal is sampled at intervals of $\Delta T, \Delta T + \Delta\tau, \Delta T + 2\Delta\tau, \Delta T + 3\Delta\tau, \ldots$; where $\Delta\tau$ p ΔT may be selected as desirable.

The sequential sampling was considered analytically and it was proven that, when the ratio $\Delta T / \Delta\tau$ is an integral number, the corresponding spectral estimates give a cut-off frequency of $\frac{1}{2\Delta\tau}$. On the contrary, when the ratio is not a whole number, the associated spectrum of the sequentially sampled data was found to comprise a complex term in its relation to the true spectrum and would not be periodic in terms of the cut-off frequency.

However, an autocorrelation function obtained from sequentially sampled data, would miss the coefficients lying in the range $R(0) \rightarrow R(\Delta T)$. These missing values could be given estimates by the initial autocorrelation coefficients estimation method, provided that the ratio $\Delta T / \Delta\tau$ is an integral number. It is because only in this case an autocorrelation matrix comprising equi-spaced coefficients would be formed and used to estimate the missing values, as seen in the Appendix.

The sequential sampling scheme would, therefore, provide means of obtaining spectral estimates with a desired cut-off frequency, when there are restrictions on the minimum allowable sampling interval.

The estimation of the autocorrelation function from sequentially sampled data was then considered and its approach was discussed. It was seen that due to the arithmetical increment of the sampling intervals, the data had to be arranged in sequences; the lag products had to be computed in each sequence and then the average values could be obtained from contributions of various sequences. The number of sample values required in each sequence was determined by the maximum desirable time delay and the ratio $\Delta T / \Delta\tau$; this was given in terms of the latter parameters. In each sequence, the lag products could be obtained from the data and specifications for this were given, which would avoid computation of the lag products corresponding to time delays exceeding the maximum required value.

It was also observed that the normalization of the auto covariance function (i.e. the division by the value at zero time delay) should take place at the end and not in each sequence.

Moreover, the digital computer simulations of the first and second order Gaussian processes, with known autocorrelation functions, were used for empirical investigations of the sequential sampling. Since this sampling scheme gives missing initial coefficients, the initial coefficients estimation method had to be applied to assign values to them. Therefore, the simulation studies could, in addition, provide some empirical considerations of the latter.

The simulation studies appeared to indicate that the accuracies of the estimated coefficients were dependent on the sample size being used. In fact, in sequential sampling, although the total sample size (used by repeating the sequence) may be relatively large but the total number of contributions to most of the lag value estimates could remain relatively small. This was because, in each sequence, the number of contributions to most of the lag value estimates could be only one or two. However, a reasonable total sample size could yield satisfactory autocorrelation estimates.

The accuracies of the estimates given to the missing initial coefficients were observed to be dependent on the accuracies of those obtained from the data. In the case of using reasonable sample sizes, the estimates given to the missing values were satisfactory; even though the estimates of those coefficients using the values assigned to other missing coefficients bore slightly less accuracies. In fact, when a small number of data sequences were used and highly inaccurate (scattered) autocorrelations were obtained, no estimates could be given to the missing values. It was noted that when this is the case, the estimating routine would notify that no values may be assigned to the missing values to maintain the corresponding autocorrelation matrix non-negative definite, as a theoretical basis for interpolation.

It is, therefore, essential to obtain reasonably accurate estimates from the data, in order to enable the assignment of the missing values. Unfortunately, the sequential sampling is a slow process and the larger the maximum time delay required for the autocorrelation function estimates, the slower would be the sampling process. Nevertheless, it was noted that in practice, for estimation of the missing coefficients, only a portion of the autocorrelation function can suffice.

In fact, when the known portion of the function is not sufficient, it would be notified by the estimating routine. However, when the missing coefficients are given estimates, then the function has been made to approximate the properties that the extrapolation is based upon and there will not be any numerical problem in extrapolating the function [18, 21]. Therefore, it was suggested that by obtaining a portion of the autocorrelation function and giving estimates to the missing values, the function may be extrapolated and used for spectral estimation. This would provide a more accurate estimation of the autocorrelation function and the spectrum with less sampling time and costs.

However, the required number of coefficients would be expected to depend on the autocorrelation function and the sampling period and does not appear to be predictable in advance. As it was pointed out, the insufficiency of this may only be found after attempting to estimate the missing values. Nonetheless, the extrapolation approach [18, 21], is employed as a basis for interpolation, i.e. in both the interpolation and extrapolation methods, the known portion of the function should be sufficient to define

its shape.

The simulation studies were then used to consider the incorporation of extrapolation method. It was observed that a reasonable total sample size (depending on the number of lag values being obtained from data) could give satisfactory autocorrelation and spectral estimates. In fact, the extrapolated values were seen to bear the imperfections of coefficients estimated from the data and the assignments made to the missing values. Some of the resulting spectra also reflected these inaccuracies and although they were not exact, they appeared to be satisfactory spectral estimates that would be obtained experimentally.

Finally, the contribution of the sequential sampling in obtaining a desired cut-off frequency was demonstrated by considering a simulation example. It was observed that in uniform sampling using the minimum allowable sampling interval could lead to an aliased spectrum and the problem was then reduced by applying the sequential sampling. The extrapolation method had also been incorporated in order to extend the autocorrelation function for spectral estimates.

APPENDIX: AN OUTLINE OF THE INITIAL AUTOCORRELATION COEFFICIENTS ESTIMATION METHOD

The estimation of the missing initial autocorrelation coefficients, i.e. when the zero lag coefficient, R(0), is known then there are a number of unknown coefficients, followed by knowledge of the coefficients describing the remaining part of the function, is based on the theoretical properties of the autocorrelation matrix and employs the same principles as the extrapolation method. Consider the autocorrelation matrix, formed by the first N coefficients, $R(0) \rightarrow R(N-1)$, of a discrete, equi-spaced, autocorrelation function, that is:

$$A_{N-1} = \begin{bmatrix} R(0) & R(1) & L & R(N-1) \\ R(1) & O & O & M \\ M & O & O & R(1) \\ R(N-1) & L & R(1) & R(0) \end{bmatrix}$$

This is a square, symmetric, non-negative definite matrix and of Toeplitz form. In fact, in the autocorrelation function extrapolation approach [18, 21], when R(N) is unknown, an estimate for this may be found such that the matrix of dimension $N+1 \times N+1$, incorporating the estimate, remains non-negative definite and has the maximum possible determinant. In this way, the autocorrelation function may be extended indefinitely in a mathematically acceptable fashion.

Now consider the case that the autocorrelation matrix comprises N+1 coefficients, $R(0) \rightarrow R(N)$, where R(1) is unknown. This unknown coefficient may be given an estimate by an iterative procedure, such that every matrix dimension greater than 2×2 is maintained non-negative definite with a maximized determinant.

The situation that several lag values, $R(1) \rightarrow R(K)$, are missing can also be considered. Let the uniform time delay of discrete autocorrelation function be $\Delta\tau$; then

R(K) should be the first coefficient to estimate by using: R(0), R(K) (to be estimated), R(2K), R(3K), ... This is equivalent to estimating a second coefficient R(1), with an effective time delay of $K\Delta\tau$. Therefore, the problem is always reducible to estimating a second coefficient R(1). Thus, the method allows the autocorrelation matrices to be formed from known coefficients, with various time delays, and give estimates to R(K), R(K-1), ..., R(1), respectively.

Furthermore, in the extrapolation routine, taking advantage of the properties of the autocorrelation matrices, being square, symmetric, non-negative definite and of Toeplitz form, there is no need to store the whole matrix and only the storage of the coefficients, R(0)→R(N), is sufficient [18, 21]. The same advantages may be equivalently taken for the iterations and computations required in the interpolation approach.

REFERENCES

[1] Aizawa, A.N. and Wah, B.W. "A Sequential Sampling Procedure for Genetic Algorithms," *Computer and Mathematics with Applications*, 27, 9-10, 77-82, 1994.

[2] Allebach, J.P. "Design of Antialiasing Patterns for Time-Sequential Sampling of Spatiotemporal Signals," *IEEE Transactions on Acoustics, Speech and Signal Processing*, 32, 1, 137- 44, 1984.

[3] Allebach, J.P. "Design of Sampling Patterns for Time Sequential Sampling of Spatio-Temporal Signals," *Proceedings of Micro-Delcon Delaware Bay Computer Conference*. 9-13, 1981.

[4] Bendat, J.S. and Piersol, A.G. *Random Data: Analysis and Measurement Procedures*. Wiley-Interscience, 1986.

[5] Bradbury, L.J.S. "Examples of the Use of the Pulsed Wire Anemometer in Highly Turbulent Flow," *Proceedings of the Dynamic Flow Conference*. Marseille, 489-509, 1978.

[6] Burrus, C.S., et. al. *Computer Based Exercises for Signal Processing Using MATLAB*. Prentice-Hall, 1994.

[7] Carlson, A. B. *Communication Systems*. McGraw-Hill, 1986.

[8] Gaster, M. and Bradbury, L.J.S. "The Measurement of the Spectra of Highly Turbulent Flows by a Randomly Triggered Pulsed-Wire Anemometer," *J. Fluid Mechanices*, 77, 499-509, 1976.

[9] Griffith, J. W. R, et. al. *Signal Processing*. Academic Press, 131-140, 1973.

[10] Jones, R. H. and Steele, N.C. *Mathematics in Communication Theory*. Ellis Horwood Publishers, 1989.

[11] Oppenheim, A.V. et. al. *Signals and Systems*. Prentice-Hall, 1983.

[12] Papoulis, A. "Maximum Entropy and Spectral Estimation: A Review," *IEEE Transaction on Acoustics, Speech, and Signal Processing*, Vol. ASSP-29, 1981.

[13] Papoulis, A. *The Fourier Integral and its Applications*. McGraw-Hill, 1962.

[14] Sarychev, V. T. "Spectral Estimation and Entropy," *Russian Physics Journal*, 36, 10, 990-9, 1993.

[15] Saucedo, R. and Schiring, E. *Introduction to Continuous and Digital Control*

Systems. Macmillan, 1968.

[16] Steward, E. G. *Fourier Optics: an Introduction.* Ellis Horwood Publishers, 1989.

[17] Taha, H. A. *Simulation Modelling and Simnet.* Prentice-Hall, 1988.

[18] Van den Bos, A. "Alternative interpretation of Maximum Entropy Spectral Analysis," *IEEE Transaction on Information Theory*, 17, 4, 493-4, 1971.

[19] Willis, N.P. and Bresler, Y. "A New Approach to the Time-Sequential Sampling Problem," *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing,* 3, 277-80, 1992.

[20] Wirth, W.D. "Energy Saving by Coherent Sequential Detection of Radar Signals with Unknown Doppler Shift," *IEE Proceedings on Radar, Sonar and Navigation,* 142, 145-52, 1995.

[21] Wu Siliang, et. al. "A New Method for Maximum Entropy Spectral Estimation," *Acta Electronica Sinica,* 22, 7, 57- 60, 1994.