Information Loss in Analog to Digital Conversion Revisited

Craig C. Smith[1], David W. Robinson[2],
David S. Hansen[1], and Bryan F. Bihlmaier[1]

[1]Department of Mechanical Engineering
Brigham Young University
Provo, UT USA
smith@byu.edu

[2]Department of Mechanical Engineering
Massachusetts Institute of Technology
Cambridge, MA USA

Abstract. When an analog signal is digitized, there is potential loss of information, even when the sample rate satisfies the Nyquist sampling criterion. This information loss can be characterized as uncertainty arising because the signal is not known before or after sampling begins, and appears as uncertainty regarding the value of the analog signal between sample points. This uncertainty is greatest near the endpoints of the sampled interval or window, and depends upon the autocorrelation or bandwidth of the signal. This error has been characterized mathematically, and guidelines have been given for reproducing analog signals from digital data to any given level of accuracy. This theory, coupled with the theory provided by Shannon and Nyquist in their well-known sampling theorem, provides a complete characterization of the information loss in sampling. Assessment of the need for simultaneous sample and hold for multiple channels in data acquisition provides a significant practical application of the theory.

1 Introduction

Analog to digital conversion has become an integral part of many systems and devices in the last couple of decades due to the development of new and effective digital technologies. Physical processes which are inherently analog are now measured, recorded, and controlled by digital processors which provide robust recordings, intelligent controllers, and great flexibility of analysis. Information processing is much more likely to be done digitally, whereas the processes and systems which the information is to characterize are often inherently analog.

The analog to digital conversion process has been characterized previously, and much is well defined. Amplitude quantization errors are well understood, and sample rate selection to avoid frequency aliasing was long ago defined by Shannon and Nyquist. What has not been defined previously is the information loss caused by truncation of the signal or frequency leakage. Although the phenomena of frequency leakage has been well understood in the context of spectrum analysis, the loss of information in terms of inability to reproduce the original analog signal within the interval over which data is sampled is the focus of the present work.

1.1 Practical Issues

In many situations, variables that are to be measured or controlled, are calculated from signals from multiple sensors. For example, temperature gradients or heat flux are found by differencing temperature sensor outputs, and system identification develops system models by relating measured inputs and outputs. The mathematical models used in these calculations usually assume that the signals used in the calculations are sampled simultaneously. Hardware for simultaneous sample and hold of multiple channels is complex and costly relative to multiplexed multiple signal samplers. But multiplexed samplers results in an offset in sample times between channels. If negligible information is lost in the process of sampling, so that the sampled, digital signal has the same information as the original analog signal, then sufficient information is present so that the original analog signal can be reproduced at any time, and the offset signals contain adequate information to characterize any desired function of the multiple signals. Thus hardware for simultaneous sample and hold is not required if care is taken to reduce all errors to a negligible level.

1.2 Categorization of Errors In Analog to Digital Conversion

Error or information loss during analog to digital conversion can be categorized into three categories: amplitude quantization error, error associated with sample rate, and error associated with signal truncation.

Amplitude Quantization Error. The conversion of the analog amplitude to a discrete amplitude is known as quantization error, and is minimized by correct matching of the corresponding ranges of the analog and digital representations, and choice of the digital word size. This error is not related to the simultaneous sample and hold question, but is here mentioned for completeness.

Errors Related to Sample Rate. The conversion from a continuous time signal to a discrete time signal (sampling) creates error known as frequency aliasing. If significant frequency components are present in the signal above one-half the sampling frequency, these components are indistinguishable from components below one-half the sampling frequency and are said to be aliased to the lower frequency. If the frequency components above one-half the folding frequency are negligible, then the information loss during sampling is negligible. This phenomenon was characterized by Shannon [1] and Nyquist [2], and is referred to as the *Shannon sampling theorem.* If significant frequency components above one-half the sampling frequency exist in the analog signal, then it is impossible to reproduce the analog signal between samples from its sampled values, even if the signal were sampled for all time.

Errors Related to Signal Truncation. Practical digital signals used for analysis or data storage must begin and end, and so some information may be lost by truncating the signal in time. If frequency domain representations of the signal are used, some assumption must always be made about the nature of the signal before data collection began and after it ended. The two most common assumptions made are (1) that the signal is periodic, and hence the data collected represents one period of a periodic function, or (2) that the signal is zero (or constant) for all time before and after data collection began.

The first of these assumptions is appropriately used for random signals that have no beginning or end, and hence cannot be assumed to be zero outside the "window" of time which is recorded. Therefore the most reasonable assumption is that the signal before and after simply repeats, but when the beginning and end of the signal are compared, it is clear that this assumption of periodicity can't be exactly correct. The error associated with this assumption of periodicity is known as frequency leakage. Since frequency components in the original signal which are not periodic in the window cannot be represented under this assumption, these components will contribute varying portions of their components to frequencies that are periodic in the window, and are said to be leaked to these other frequencies.

In transient signals for which the second assumption is valid, the signal is known for all time (it is zero everywhere outside of where it is recorded), and no loss of information occurs during truncation, as long as the truncation is done without artificially introducing high frequencies (above one-half the sampling frequency) during truncation. If a transient signal is truncated before it ends, so that it is not effectively zero outside the interval over which it is recorded, then frequency leakage represents the loss of information associated with this truncation.

Clearly frequency leakage will always occur whenever we don't know what a signal is before and after (outside of the window) we recorded it. The only way to recover this information loss is to find out what we didn't record. The most effective way to minimize this loss of information is to record the signal for a longer period of time. Since it is not practical to record any signal forever, and computer memory, data storage, and analysis requirements grow as data array sizes grow, any data collection system must be designed to minimize loss of information by truncation or frequency leakage by trading off the length of time data will be recorded with other issues and constraints.

Summary of A/D Conversion Error. A data collection and analysis system is designed so that loss of information during analog to digital conversion is negligible by

1) matching the level of the analog signal and converter range, and choosing the digital word size so amplitude quantization errors are negligible,
2) sampling at a frequency of at least twice the highest non-negligible frequency component in the signal so that frequency aliasing is negligible, and
3) collecting data for a long enough time so that frequency leakage errors are negligible.

If these requirements are met, then sufficient information exists in the sampled (digital) signal to reconstruct the information content of the original analog signal. Hence, information is present to determine the value of the original signal at any point in time required, and the time of sampling is irrelevant to the information content. Therefore, the only justification for simultaneous sample and hold hardware is to simplify or speed the processing in such cases as this is necessary. Otherwise, sampling offset between two signals can always be compensated by processing with negligible loss of information, making simultaneous sample and hold or synchronized separate A/D converters less cost effective than multiplexing systems.

2 Quantitative Description of Information Loss in Signal Truncation

Given a digital sequence, x_k, sampled without frequency aliasing at sample interval, T, the signal can be reproduced between samples by filtering with an "ideal" filter that cuts off exactly at the folding frequency, $f_f = \frac{1}{2T}$. This is accomplished mathematically by convolution with the inverse transform of the frequency response function, which is the sinc function [3]:

$$\text{sinc}\; pt = \frac{\sin pt}{pt} \qquad (1)$$

where non-dimensional time, $\tau = t/T$, has been used for convenience. The resulting continuous signal, $y(\tau)$ can thus be represented from the digital signal, x_k, by the relation

$$y(t) = \sum_{k=-\infty}^{\infty} x_k \,\text{sinc}\,[p(t-k)] = \sum_{k=0}^{N-1} x_k \,\text{sinc}\,[p(t-k)] + \sum_{k=-\infty}^{-1} x_k \,\text{sinc}\,[p(t-k)] + \sum_{k=N}^{\infty} x_k \,\text{sinc}\,[p(t-k)] \qquad (2)$$

Since x_k is only known over the interval $0 = k = N-1$, only the first summation can be evaluated directly. If x_k is periodic, the above convolution is easily accomplished in the frequency domain, since a Discrete Fourier Transform (DFT) can be found using the Fast Fourier Transform algorithm, and the convolution process becomes multiplication in the frequency domain. If x_k is only known for the interval k = 0 through N-1, error results in either neglecting the terms for which x_k is unknown (effectively assuming $x_k = 0$ outside this range), or in assuming the periodic form p_k, where p_k is the discrete signal that is identical to x_k over the interval k=0 through N-1, and then repeats at intervals of N data points. Letting $\tilde{y}(\tau)$ be the result of doing the convolution with the sequence p_k, the difference

$$\begin{aligned}
[y(t) - \tilde{y}(t)] &= \sum_{k=0}^{N-1} x_k \,\text{sinc}\,[p(t-k)] - \sum_{k=0}^{N-1} p_k \,\text{sinc}\,[p(t-k)] + \sum_{k=-\infty}^{-1} x_k \,\text{sinc}\,[p(t-k)] - \sum_{k=-\infty}^{-1} p_k \,\text{sinc}\,[p(t-k)] \\
&+ \sum_{k=N}^{\infty} x_k \,\text{sinc}\,[p(t-k)] - \sum_{k=N}^{\infty} p_k \,\text{sinc}\,[p(t-k)] \\
&= \sum_{k=0}^{N-1} (x_k - p_k)\,\text{sinc}\,[p(t-k)] + \sum_{k=-\infty}^{-1} (x_k - p_k)\,\text{sinc}\,[p(t-k)] + \sum_{k=N}^{\infty} (x_k - p_k)\,\text{sinc}\,[p(t-k)]
\end{aligned} \qquad (3)$$

But since $p_k = x_k$ for k=0 through N-1, the first summation vanishes and

$$[y(t) - \tilde{y}(t)] = \sum_{k=-\infty}^{-1} (x_k - p_k)\,\text{sinc}\,[p(t-k)] + \sum_{k=N}^{\infty} (x_k - p_k)\,\text{sinc}\,[p(t-k)] \qquad (4)$$

If $\tau << N$, the terms in the second summation and terms for $k < -N$ in the first summation are negligible, and

$$[y(t) - \tilde{y}(t)] = \sum_{k=-N}^{-1} (x_k - p_k)\,\text{sinc}\,[p(t-k)] \qquad (5)$$

Squaring both sides of the equation above, we can write

$$
\begin{aligned}
[y(t) - \tilde{y}(t)]^2 &\approx \left[\sum_{k=-N}^{-1} (x_k - p_k)\,\mathrm{sinc}\,[p(t-k)]\right]\left[\sum_{r=-N}^{-1} (x_r - p_r)\,\mathrm{sinc}\,[p(t-r)]\right] \\
&= \sum_{k=-N}^{-1}\sum_{r=-N}^{-1} (x_k - p_k)(x_r - p_r)\,\mathrm{sinc}[p(t-k)]\,\mathrm{sinc}[p(t-r)] \\
&= \sum_{k=-N}^{-1}\sum_{r=-N}^{-1} (x_k x_r - p_k x_r - x_k p_r + p_k p_r)\,\mathrm{sinc}[p(t-k)]\,\mathrm{sinc}[p(t-r)] \\
&= \sum_{k=-N}^{-1}\sum_{r=-N}^{-1} x_k x_r\,\mathrm{sinc}[p(t-k)]\,\mathrm{sinc}\,[p(t-r)] - \sum_{k=-N}^{-1}\sum_{r=-N}^{-1} p_k x_r\,\mathrm{sinc}[p(t-k)]\,\mathrm{sinc}\,[p(t-r)] \\
&\quad - \sum_{k=-N}^{-1}\sum_{r=-N}^{-1} x_k p_r\,\mathrm{sinc}[p(t-k)]\,\mathrm{sinc}\,[p(t-r)] + \sum_{k=-N}^{-1}\sum_{r=-N}^{-1} p_k p_r\,\mathrm{sinc}[p(t-k)]\,\mathrm{sinc}\,[p(t-r)]
\end{aligned}
\tag{6}
$$

Since p_k is periodic, and is equal to x_k for 0 = k = N-1, over the interval -N = k = -1, $p_k = x_{k+N}$, thus:

$$
\begin{aligned}
[y(t) - \tilde{y}(t)]^2 &\approx \sum_{k=-N}^{-1}\sum_{r=-N}^{-1} x_k x_r\,\mathrm{sinc}[p(t-k)]\,\mathrm{sinc}\,[p(t-r)] - \sum_{k=-N}^{-1}\sum_{r=-N}^{-1} x_{k+N} x_r\,\mathrm{sinc}[p(t-k)]\,\mathrm{sinc}\,[p(t-r)] \\
&\quad - \sum_{k=-N}^{-1}\sum_{r=-N}^{-1} x_k x_{r+N}\,\mathrm{sinc}[p(t-k)]\,\mathrm{sinc}\,[p(t-r)] + \sum_{k=-N}^{-1}\sum_{r=-N}^{-1} p_k p_r\,\mathrm{sinc}[p(t-k)]\,\mathrm{sinc}\,[p(t-r)]
\end{aligned}
\tag{7}
$$

Averaging both sides of the equation, we substitute the autocorrelation functions, and for convenience, we convert the summation range to positive integers by changing signs on the subscripts k and r:

$$
\begin{aligned}
E_p\left[[y(t) - \tilde{y}(t)]^2\right] &\approx \sum_{k=1}^{N}\sum_{r=1}^{N} R_x(r-k)\,\mathrm{sinc}[p(t+k)]\,\mathrm{sinc}[p(t+r)] - \sum_{k=1}^{N}\sum_{r=1}^{N} R_x(r-k+N)\,\mathrm{sinc}[p(t+k)]\,\mathrm{sinc}[p(t+r)] \\
&\quad - \sum_{k=1}^{N}\sum_{r=1}^{N} R_x(r-k-N)\,\mathrm{sinc}[p(t+k)]\,\mathrm{sinc}[p(t+r)] + \sum_{k=1}^{N}\sum_{r=1}^{N} R_p(r-k)\,\mathrm{sinc}[p(t+k)]\,\mathrm{sinc}[p(t+r)]
\end{aligned}
\tag{8}
$$

where $R_x(k)$ and $R_p(k)$ are the autocorrelation functions of x_k and p_k respectively, and the operator $E_p[\ \ ]$ denotes the expected value using the periodic function p_k to estimate x_k outside the interval over which the signal is recorded.

Without loss of generality, we now assume that the signal, x(t), has zero mean, so that the autocorrelation function $R_x(t)$ represents the variance, and $R_x(t) \longrightarrow 0$ as t ⟶ 8 . Further, we will assume that the discrete form, $R_x(q) = 0$ for $q = -\frac{N}{2}$ and for $q > \frac{N}{2}$. Using this assumption, and also noting that $R_p(q \pm pN) = R_x(q)$ for all integers, p (i.e. $R_p(q)$ is periodic), some algebraic manipulation results in the following :

$$
\begin{aligned}
E_p\left[[y(t) - \tilde{y}(t)]^2\right] &\approx 2\sum_{k=1}^{\frac{N}{2}}\ \sum_{r=1}^{\frac{N}{2}+k} R_x(r-k)\,\mathrm{sinc}\,[p(t+k)]\ \mathrm{sinc}\,[p(t+r)] \\
&\quad + 2\sum_{k=\frac{N}{2}+1}^{N}\ \sum_{r=k+1-\frac{N}{2}}^{N} R_x(r-k)\ \mathrm{sinc}\,[p(t+k)]\ \mathrm{sinc}\,[p(t+r)]
\end{aligned}
\tag{9}
$$

Equation (9) represents the mean-squared error in reproducing the analog signal from the digital signal under the assumption that the signal is periodic (the collected data is one period of a periodic function). Although this assumption is known to be incorrect, the economy of using the Fast Fourier Transform (FFT) to reproduce the signal at desired points in time makes this assumption desirable. If the above process is repeated with the assumption that the signal is zero[1] outside the interval over which the data is collected, the mean squared error is reduced to exactly one-half the value under the periodic assumption:

$$
E_o\left[[y(t) - \tilde{y}(t)]^2\right] \approx \frac{1}{2}\,E_p\left[[y(t) - \tilde{y}(t)]^2\right] \tag{10}
$$

Although this method reduces the mean squared error by a factor of two, it requires zero padding when using the FFT to determine signal values between samples, effectively increasing the FFT array size by a factor of two for a given data set.

3. Experimental Verification

[1] If the signal has non-zero mean, this assumption is equivalent to assuming that the signal is equal to the mean value of the signal in the interval over which it is collected.

To verify the analytical description above, pseudo analog signals, sampled at various times were created. From signals sampled at one time, signals sampled at other times were recreated. The average errors in recreating these signals at alternate sample times were compared with the analytical error predictions above.

3.1 Pseudo Analog Signal

A series of normally distributed random numbers was generated, representing a pseudo analog signal. This "analog" signal included enough numbers to ensure no transients were introduced from the digital filter, as well as to account for values of the signal at intermediate "time" steps. For example, suppose a block of data with eight data points was desired, with a desired known value of the signal one-half time step forward of each point. A random number sequence, X, is created using the random number generator,

$$X = \{x_1, x_2, x_3, \ldots, x_{14}, x_{15}, x_{16}\}$$

where x_1, x_3, ..., represent the analog signal at each time step, and x_2, x_4, ..., represent the value of the analog signal at intervals half-way between the time steps.

Next, the signal was digitally filtered using an 8th-order butterworth filter with a cut-off frequency of one-half the folding frequency (one-fourth the sampling frequency) -- well inside the Nyquist criterion. The characteristics of the butterworth filter can be observed in the unshifted signal's autospectrum, Figure 1.

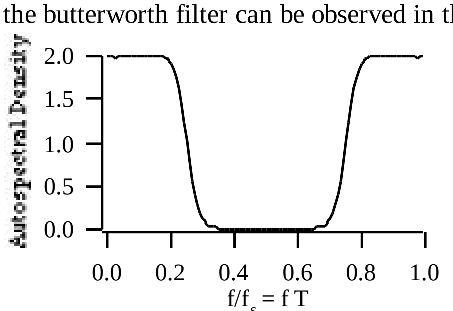

Figure 1. Autospectrum of Unshifted Pseudo Analog Signal

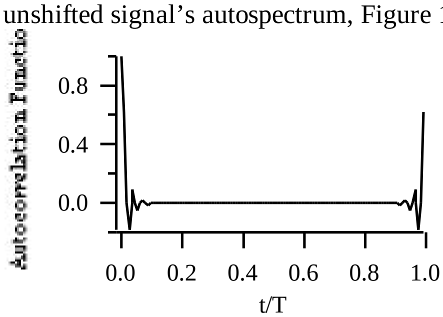

Figure 2. Autocorrelation of Unshifted Pseudo Analog Signal

The filtering at one-half the folding frequency was to ensure no aliasing and to define the signal's autocorrelation function, Figure 2. Scaling of the filtered signal ensured a statistical mean of zero and mean square of one. These filtered random data now represented a pseudo analog signal, with known bandwidth and frequency content.

The filtered signal was then stored in two vectors. The first vector represented a sampled signal which we desire to shift in time. The second vector represented the exact value expected of the first signal following the time shift. Following the above example,

$$X1 = \{x_1', x_3', x_5', \ldots, x_{15}'\} \qquad \text{and} \quad X2 = \{x_2', x_4', x_6', \ldots, x_{16}'\}$$

where the prime (') denotes filtered data. These two new sequences were thus two versions of the same band-limited signal sampled at offset time sequences.

The first signal was then shifted forward in time, using the process described below. Finally, the mean square error between the shifted (first) signal and the exact (second) signal was calculated. This entire process was accomplished for 19 intermediate times between whole time steps.

This process was repeated and the results were averaged to account for the faults in the random number generation. It will be noted that even with 500,000 averages, the assumed white noise of the unshifted signal is not completely "flat" in the lower frequency range, although this is not noticeable because of the scale in Figure 1 plotted at small size to save space. This higher frequency content can be observed as well in the non zero values in the midrange of the autocorrelation of the unshifted signal, again not noticeable at the scale of Figure 2.

3.2 Time Shift of a Digital Signal

The Fourier transform is used to map continuous signals from the time domain to the frequency domain. Defining h(t) and H(f) as a Fourier transform pair:

$$h(t) \Leftrightarrow H(f)$$

An original time signal shifted in time by an offset t transforms as a frequency signal multiplied by a phase shift.

$$h(t - t) \Leftrightarrow H(f)e^{j2pft}$$

Similar to the continuous domain above, a block of discrete time data is transformed to the frequency domain using the discrete Fourier transform. Additionally, if the block size of the data is of the power 2^n, where n is an integer, the fast Fourier transform (FFT) can be used.

$$h(k) \Leftrightarrow H(n)$$

To shift in time, a phase shift is again applied in the frequency domain. Notice the modification of the phase shift for use with discrete data, but also its similarity in form to that of the continuous transform above.

$$h(k - t) \Leftrightarrow H(n)e^{-j\frac{2pnt}{N}}$$

where k = time sample number, n = frequency sample number, N = total number of data points in the data block, and r = the offset between data points (-1 < r < 1) representing the fraction of the sample interval for the shift between data sequences. The inverse FFT of the phase-shifted data results in the time-shifted discrete time signal.

3.3 Comparison to Analytical Expression

The values of the mean square error calculated from the above process were compared to the mean square error calculated from the analytical expression. To calculate a mean square error from the analytical expression, it was necessary to determine the characteristics of the digital butterworth filter. As broad band random noise passes through a butterworth filter, it essentially takes on the characteristic shape of the filter, scaled by the magnitude of the noise. Given the characteristic of the filter, the autocorrelation of a broad band random signal can be determined analytically.

The numerator and denominator of the characteristic polynomial (z transform) for the eighth-order butterworth filter were supplied as input to the "freqz" command in Matlab, giving the complex frequency response function of the butterworth filter. This frequency response function was then multiplied by its complex conjugate to form the autospectrum of the periodic filtered signal, since the input to the filter was white noise. The inverse Fourier transform of the autospectrum, the autocorrelation, was then scaled such that the mean square of the autocorrelation was equal to one. This autocorrelation was then used in the analytical expression for R_X, which goes to zero at the data midpoint and remains at zero.

Comparisons of the mean squared error from the digital experiment to the analytical expression (Equation 9) over the interval from -1 = t = 3 (beginning one time step before data is collected) are shown below in Figure 3. The solid line represents the analytical expression, while the asterisks represent the experimental data. As can be seen, agreement is excellent.

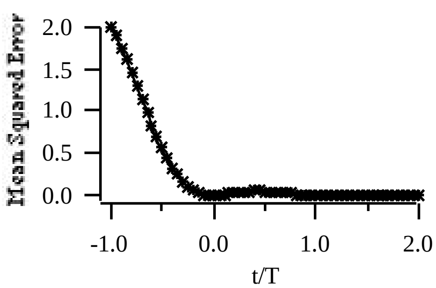

Figure 3a Comparison of analytical expression to experimental results

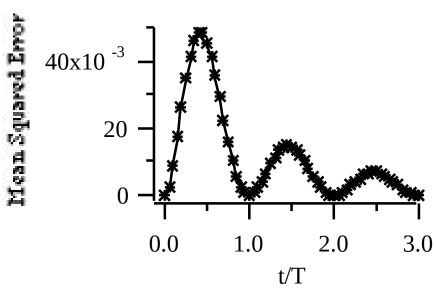

Figure 3b. Comparison of analytical expression to experimental results

Although the analytical expression was developed for low values of time (near zero), by symmetry the error expression should be valid if time is interpreted as measured from either the beginning (forward) or the end (backward) of a dataset. This was also confirmed by comparison with the experimental data described above, and yielded results similar to those shown in Figure 3.

4 Effect of Bandwidth on Root Mean-Squared Error

Since the mean squared error in reproducing a signal between sample points is a function of the autocorrelation function, which relates directly to the bandwidth of the signal, it is instructive to examine the sensitivity of this error to the bandwidth of the signal. Although practical signals rarely (if ever) have flat spectra over some band of frequencies with no frequency content outside that band, such spectra represent a limiting case in that they represent signals which are correlated over somewhat longer intervals of time than other signals of the same essential bandwidth but with smoother drop in spectra with frequency. Thus we will use these "ideal bandpass" signals to compare the effects of bandwidth. The autocorrelation for such a signal of bandwidth $f_b = bf_f$, where f_f is the folding frequency (one-half the sampling frequency) can be written as

$$R_b(k) = \mathrm{sinc}^2(bpk) = \frac{\sin^2(bpk)}{(bpk)^2} \qquad (11)$$

The case where b = 1 represents the maximum bandwidth possible for a digital signal, and is therefore a limiting case (maximum error possible). For this case, $R_1(k) = 1$ for k=0 and $R_1(k) = 0$ for k?0, and Equation (9) reduces to

$$E_1\left[[y(t) - \tilde{y}(t)]^2\right] = 2\sum_{n=1}^{N} \mathrm{sinc}^2\,[p(t+n)] \qquad (12)$$

and the root mean squared error is found by taking the square root. Using various values of bandwidth fraction, b, the root mean squared error for various bandwidths are shown in Figure 4. It is noted that as the bandwidth fraction, b, gets smaller, the error fairly quickly approaches the limiting case where b = 0. For this limiting case, $R_0(k) = 1$ for all k (perfect correlation of the signal with itself for all times).

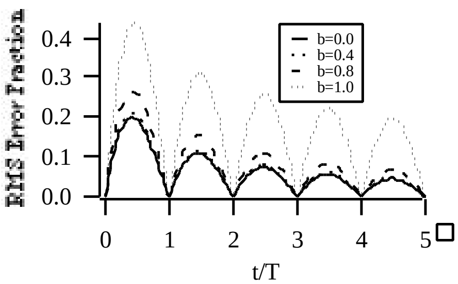

Figure 4. Root mean squared error for various signal bandwidth

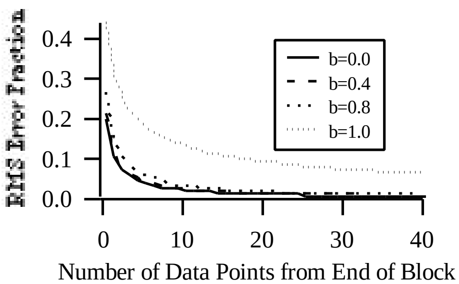

Figure 5. The maximum rms error in reproducing an analog signal within a given number of points from the end.

The maximum error in reproducing the signal at any time within a given number of data points from the end is shown in Figure 5. This illustrates even more graphically how quickly the error bounds approach the limit as b approaches 0. Practically, real signals will not have bandwidths near the folding frequency to avoid aliasing. Anti-aliasing filters used in data acquisition will typically cutoff at b ˜ 0.5 to 0.8. Since typical real signals will be band limited before filtering, it would be expected that b = 0.4 would be representative of many practical signals.

5 Practical Reconstruction of Analog Signals at Alternate Sample Times

Analog signals can thus be reconstructed between samples, with error depending upon the time from the end of the collected samples and the position between samples of the time to reconstruct. To perform the reconstruction requires some assumption about the nature of the signal outside the window of time for which it is collected. When two signals are recorded via a multiplexer, the samples for one are offset from the samples of the other by some fraction of the sample interval, depending upon how many signals are multiplexed and the configuration of the multiplexer. The most efficient way computationally to reconstruct one signal at sample times simultaneous with the other is to use the FFT algorithm to transform one time series to the frequency domain. Each frequency term is then shifted by the phase corresponding to the multiplexer offset (phase shift is proportional to frequency). The resulting signal at the offset sample times is then found by using the inverse FFT. Since the FFT implicitly assumes the signal to be periodic (this is required because the frequency spectrum is discrete), the natural assumption is to assume that the signal is periodic in the window. The assumption that the signal is constant outside the window can also be implemented computationally using the FFT by padding the time series with the mean value of the signal, doubling the size of the array to be transformed making at least half of the terms in the array equal to the signal mean. This doubling of the array size reduces the mean squared error to one-half that of the error resulting without padding, where the assumption is that the signal is periodic in the window. Since the error decreases with the time from the end of the array, to reduce the error to the desired level, additional data points should be collected beyond the region for which the data samples will need to be accurately shifted, and then data at the ends of the interval can be discarded. It is expected that for many cases, the number of additional data points that would have to be discarded under the periodic assumption (vs. the constant value assumption) would be less than the number of data points required to perform the padding. Therefore, the periodic assumption is the most practical, even though it increases the mean squared error by a factor of 2 (and rms error by a factor of $\sqrt{2}$ as compared to the constant value assumption.

Table 1 Number of Data Points at Each End of a Time Series With Error Above the Specified Level After a Time Shift

$f_b = f_f$	RMS Error Ceiling:				
Forward-shift Time-fraction	5%	2.5%	1%	0.5%	0.1%
0.1	7	26	178	599	*
0.2	23	106	551	1580	*
0.3	48	194	916	3665	*
0.4	65	260	1167	4669	*
0.5	71	284	1256	5024	*

$f_b = 0.8f_f$	RMS Error Ceiling:				
Forward-shift Time-fraction	5%	2.5%	1%	0.5%	0.1%
0.1	2	4	10	21	100
0.2	4	8	19	40	177
0.3	5	11	26	54	232
0.4	6	12	32	63	266
0.5	6	13	34	66	277

$f_b = 0.4f_f$	RMS Error Ceiling:				
Forward-shift Time-fraction	5%	2.5%	1%	0.5%	0.1%
0.1	1	3	7	17	77
0.2	3	5	14	28	153
0.3	4	8	20	39	218
0.4	4	9	23	46	261
0.5	4	9	24	49	276

$f_b = 0$ (limiting case)	RMS Error Ceiling:				
Forward-shift Time-fraction	5%	2.5%	1%	0.5%	0.1%
0.1	1	2	7	14	65
0.2	2	5	12	26	116
0.3	3	7	17	35	152
0.4	4	8	19	41	174
0.5	4	8	20	43	181

* Number of data points required is outside useful range of data collection.

To aid in determining the number of data points that would need to be dropped from each end of the time series to obtain a given level of error after a time shift, Table 1 has been developed. If data is to be reproduced at times other than the original sample times, one first determines the desired accuracy for which the data must be reproduced. Then for the desired offset between the time at reconstruction and the original sample times, and for the given bandwidth of the signal, the number of data points that will have error beyond the desired level is determined. A data block is selected such that the total number of data points includes the desired extra data points at each end of the data to be retained after reconstruction. Overlapping blocks can thus be used to reproduce accurately a long time series at alternate sample times.

6 Conclusions

When an analog signal is digitized, there is potential loss of information, even when the sample rate satisfies the Nyquist sampling criterion. This information loss can be characterized as uncertainty arising because the signal is not known before or after sampling begins, and appears as uncertainty regarding the value of the analog signal between sample points. An analytical expression has been developed to describe this uncertainty as a function of time after sampling of the data begins or (by symmetry) as a function of time before sampling of the data ends. The analytical expression has been verified by numerical simulation. This uncertainty is greatest near the endpoints of the sampled interval or window, and depends upon the autocorrelation or bandwidth of the signal. Guidelines have been given for reproducing analog signals from digital data to any given level of accuracy. This theory, coupled with the theory provided by Shannon and Nyquist in their well-known sampling theorem, provides a complete characterization of the information loss in sampling analog data. Assessment of the need for simultaneous sample and hold for multiple channels in data acquisition provides a significant practical application of the theory.

References

[1] Shannon, C.E., "Communication in the Presence of Noise," *Proceedings of the IRE*, pp.10-21, January, 1949

[2] Nyquist, H., "Certain Factors Affecting Telegraph Speed," *Bell Systems Technical Journal*, Vol 3, p. 324, April, 1924

[3] Proakis, J.G., and Manolakis, D.G., *Digital Signal Processing Principles, Algorithms, and Applications*, Second Edition, Macmillan Publishing Co., 1992