

On Error-Rate Characteristics of Oversampled A/D Conversion

Zoran Cvetković and Martin Vetterli

Abstract— Accuracy of simple A/D conversion¹ can be improved by refining either sampling interval or quantization step. Although the former is more convenient for implementation reasons, the latter seems to be more efficient. When the quantization step tends to zero for a fixed sampling interval, the expected value of the quantization error decays exponentially as a function of the bit-rate B , $E(e(t)^2) = O(\exp(-2\alpha B))$. On the other hand, if the oversampling ratio is increased for a fixed quantization step the error behaves as $E(e(t)^2) = O(1/B)$. We show that this inferior error-rate performance of oversampled A/D conversion is a consequence of inadequate reconstruction and encoding, and propose an efficient scheme for lossless encoding of quantized samples. With this encoding scheme the error of oversampled A/D conversion decays exponentially with the bit-rate.

I. INTRODUCTION

Analog to digital (A/D) conversion requires discretization of an analog signal in both time and amplitude. Discretization in time is implemented as sampling with a sampling interval τ , while the discretization in amplitude amounts to quantization of these samples with a quantization step q . The commonly used linear reconstruction restores the analog signal from the digitally encoded representation by low-pass filtering a pulse train modulated by the sequence of quantized samples (see Figure 1). If the considered analog signal $f(t)$ is bandlimited and τ is smaller than the Nyquist sampling interval, τ_N , the discretization in time is a reversible operation. However, the amplitude discretization introduces an irreversible loss of information. Hence, a reconstructed signal $f_r(t)$ is generally different from the original $f(t)$. The error $e(t) = f_r(t) - f(t)$ is denoted as *quantization error*.

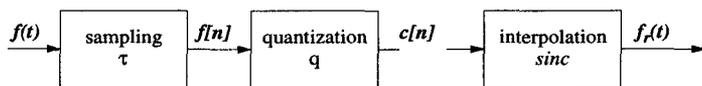


Fig. 1. Block diagram of simple A/D conversion followed by classical reconstruction. Input σ -bandlimited signal $f(t)$ is first sampled at a frequency $f_s = 1/\tau$, which is above the Nyquist frequency. The sequence of samples $f[n]$ is then discretized in amplitude with a quantization step q . Classical reconstruction gives a signal $f_r(t)$, which is obtained as a low-pass filtered version of some signal having the same digital version as the original $f(t)$.

A common way to study the quantization error is statistical analysis that models $e(t)$ as a uniformly distributed

The authors are with École Polytechnique Fédérale de Lausanne, Dépt. d'Electricité, LCAV, CH-1015 Lausanne, Switzerland. M. Vetterli is also with the University of California at Berkeley, Berkeley, CA 94720, USA. E-mail: Zoran.Cvetkovic@de.epfl.ch, Martin.Vetterli@de.epfl.ch

¹Considerations in this paper are restricted to simple A/D conversion as opposed to $\Sigma\Delta$ modulation.

white noise independent of the input. Assuming the white noise model, one can show that the mean squared error is given by

$$E(e(t)^2) = \frac{1}{12} q^2 \frac{\tau}{\tau_N}, \quad (1)$$

which agrees with measured data under conditions such as having a large number of quantization levels and a small quantization step compared to the input amplitude range [1]. This result demonstrates that conversion accuracy can be improved by refining either quantization or sampling. Since the quantization step size is constrained by the circuit complexity and the precision of analog components, high accuracy of modern A/D converters is usually based on oversampling. However, if we consider the error as a function of the bit-rate, quantization refinement outperforms oversampling. As an illustration, consider the case when the quantization step is halved. This reduces the mean squared error by a factor of four, at the expense of an additional bit per sample. In order to achieve the same error reduction by refining the sampling interval, the oversampling ratio $R = \tau_N/\tau$, should be increased four times, and consequently the total number of bits increases by the same factor.

This reasoning gives the following results. If the quantization step is decreased for a fixed sampling interval τ , the error decays exponentially with the bit-rate B . Precisely,

$$E(e(t)^2) = \frac{1}{12} A^2 \frac{\tau}{\tau_N} 2^{-2\tau B}, \quad (2)$$

where A denotes signal amplitude range. On the other hand if the sampling is successively refined, $\tau \rightarrow 0$, for a fixed quantization step q ,

$$E(e(t)^2) = \frac{1}{12} \frac{q^2}{\tau_N} \left(\log_2 \frac{A}{q} \right) \frac{1}{B}, \quad (3)$$

or in other words the mean squared error decays only linearly with the bit-rate.

These results are obtained with linear reconstruction (low-pass filtering) and a pulse code modulation (PCM) scheme for lossless encoding of the quantized samples. The point of this paper is to demonstrate that the inferior performance of oversampling with respect to quantization refinement is a result of inadequate reconstruction and coding. We show that if an efficient encoding scheme is used together with an optimal reconstruction, the quantization error decays in the squared norm as

$$\|e(t)\|^2 = O(2^{-2\beta B}), \quad (4)$$

when $\tau \rightarrow 0$ for a fixed quantization step.² The factor β in (4) is of the same order of magnitude as the Nyquist sampling interval, so that we have performance comparable to that when the quantization step tends to zero, and significant improvement compared to traditional results.

Notation

In this paper, the term bandlimited signals will refer to both periodic bandlimited signals and σ -bandlimited signals, unless otherwise stated. A signal $f(t)$ is said to be σ -bandlimited if it is square integrable, $f(t) \in L^2(\mathbf{R})$, and its Fourier transform, $\hat{f}(\omega)$, satisfies $|\hat{f}(\omega)| = 0, |\omega| > \sigma$. The norm of the σ -bandlimited signal $f(t)$ is given by

$$\|f(t)\|^2 = \int_{-\infty}^{\infty} |f(t)|^2 dt.$$

A space of periodic bandlimited signals, with a period T , is the space of trigonometric polynomials

$$f(t) = a_0 + \sum_{i=1}^N (a_i \cos(2\pi i t/T) + b_i \sin(2\pi i t/T)).$$

The dimension of the space of periodic bandlimited signals given by this expression is $2N + 1$. The norm of a signal $f(t)$ in this space is given by

$$\|f(t)\|^2 = \int_0^T |f(t)|^2 dt.$$

The Nyquist sampling interval, τ_N , in the space of periodic bandlimited signals with a period T and dimension $2N + 1$ is equal to $\tau_N = T/(2N + 1)$. In the space of σ -bandlimited signals it is given by $\tau_N = \sigma/\pi$.

II. ERROR REDUCTION WITH CONSISTENT RECONSTRUCTION

Information contained in the digital representation, generated as the result of A/D conversion, allows for reconstruction of the input analog signal with an error $e(t)$ which behaves in the squared norm as $\|e(t)\|^2 = O(\tau^2)$. This fact was first pointed out for the case of periodic bandlimited signals [2], and later also proven for bandlimited signals in $L^2(\mathbf{R})$ [3]. Recall that the traditional result, which pertains to linear reconstruction, asserts that the mean squared quantization error behaves as $E(e(t))^2 = O(\tau)$. The improved performance can be achieved with so called *consistent reconstruction*. This term denotes an algorithm which always gives a reconstructed bandlimited signal which after A/D conversion produces the digital sequence which is identical to the sequence produced by the original. In the context of consistent reconstruction, quantization error is studied using deterministic analysis [2], [3], and for that reason error bounds in this section are given with respect to its squared norm rather than its variance. Note that in the case of sampling at the Nyquist

²Please see the notation part below for the definition of the norm $\|\cdot\|$.

rate, linear reconstruction is also consistent, however this need not be true if some oversampling is introduced. Algorithms for consistent reconstruction were proposed in [2], [5], and were based on alternating projections onto sets of convex constraints defined by the digital representation.

In order to achieve reconstruction of a bandlimited signal $f(t)$ with an error $e(t)$ bounded as

$$\|e(t)\|^2 \leq K\tau^2, \quad (5)$$

it is essential that samples taken at points of quantization threshold crossings of $f(t)$ provide complete and stable description of the corresponding class of bandlimited signals [2], [3], [4]. If a space of periodic bandlimited signals is considered, this condition is satisfied by any set of points as long as its cardinality is greater than or equal to the dimension of the space. The constant of proportionality K in (5) depends on the norm of $f(t)$ in the respective space, and also on the distribution of its quantization threshold crossings [2], [3]. It is important to note that if $f(t)$ is a periodic bandlimited signal, its digital representation allows for reconstruction with an error which either tends to zero with increased oversampling as $\|e(t)\|^2 = O(\tau^2)$, or does not approach zero at all. The latter occurs when $f(t)$ does not have a sufficient number of quantization threshold crossings. The situation with σ -bandlimited signals is much more subtle [3]. However, if the reconstruction error is to converge to zero as $\tau \rightarrow 0$, then the samples taken at quantization threshold crossings of the considered signal should provide a complete and stable description of signals in the given space of σ -bandlimited functions, and in that case $\|e(t)\|^2 = O(\tau^2)$.

III. AN EFFICIENT ENCODING SCHEME

An efficient scheme for lossless encoding of the digital representation follows from the observation that for sufficiently fine sampling the quantized values of signal samples can be determined from the corresponding sequence of quantization threshold crossings. Consider a bandlimited

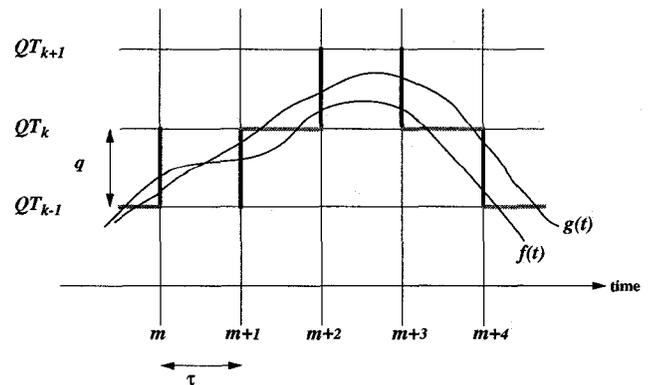


Fig. 2. Quantization threshold crossings. Two bandlimited signals $f(t)$ and $g(t)$, having the same pattern of quantization threshold crossings, also have the same quantized values of corresponding samples.

signal $f(t)$ and suppose that its quantization threshold

crossings are separated, that is there is an $\epsilon > 0$ such that no two quantization threshold crossings are closer than ϵ . Note that a bandlimited signal of a finite energy $f(t)$ has a bounded slope so there is always an $\epsilon_1 > 0$ such that $f(t)$ can not go through more than one quantization threshold on any interval shorter than ϵ_1 . The condition for separated quantization threshold crossings requires in addition that intervals between consecutive crossings through the same threshold are limited from below away from zero. For a sampling interval smaller than ϵ all quantization threshold crossings of $f(t)$ occur in distinct sampling intervals. Under this condition, quantized samples of $f(t)$ are completely determined by the sequence of quantization threshold crossings of $f(t)$ (see Figure 2). Another effect of high oversampling is that quantized values of consecutive samples differ with a small probability. In this case, an economical digital representation would encode incidences of data changes, that is sampling intervals where quantization threshold crossings occur, rather than quantized samples themselves. One such encoding scheme is described below.

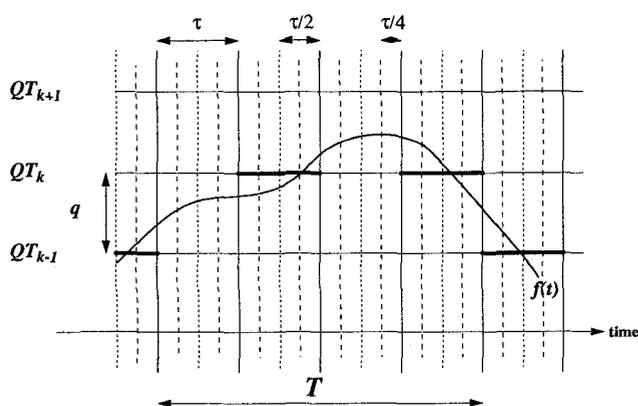


Fig. 3. Quantization threshold crossings encoding. Quantization threshold crossings are grouped on intervals of a length T . Refining the sampling interval by a factor 2^k requires additional k bits per quantization threshold crossing to encode its position inside the interval T .

Quantization threshold crossings can be grouped on consecutive time intervals of a given length T . For each of the crossings at most $1 + \log_2(T/\tau)$ bits are then needed to record its position inside the interval T . The height of the threshold crossing can be given with respect to the previous one, so that for this information only one additional bit is needed, to denote the direction of the crossing (upwards or downwards). Hence, for recording of the information on quantization threshold crossings on an interval where Q of them occur, less than $Q(2 + \log_2(T/\tau))$ bits are needed (see Figure 3). The bit-rate is then bounded as

$$B \leq \frac{Q_m}{T} \left(2 + \log_2 \left(\frac{T}{\tau} \right) \right), \quad (6)$$

where Q_m denotes the maximal number of the crossings on an interval of length T . If this efficient coding is used together with consistent reconstruction, then it follows from

(5) and (6) that the error-rate characteristics of the over-sampled A/D conversion become

$$\|e(t)\|^2 = 16KT^2 2^{-2} \frac{T}{Q_m} B, \quad (7)$$

or $\|e(t)\|^2 = O(\exp(-2\beta B))$, where $\beta = T/Q_m$.

IV. AN ESTIMATE OF FACTOR β

Considerations in this subsection pertain to bandlimited signals in $L^2(\mathbf{R})$, since corresponding derivations for the case of periodic bandlimited signals are straightforward. In order to estimate the factor T/Q_m in (7), we consider again the two types of quantization threshold crossings, denoting them as d -crossings and s -crossings, with the following meanings. A quantization threshold crossing is said to be a d -crossing if it is preceded by a crossing of a different quantization threshold, or s -crossing if it is preceded by a crossing of the same threshold (see Figure 4). The total number of quantization threshold crossings

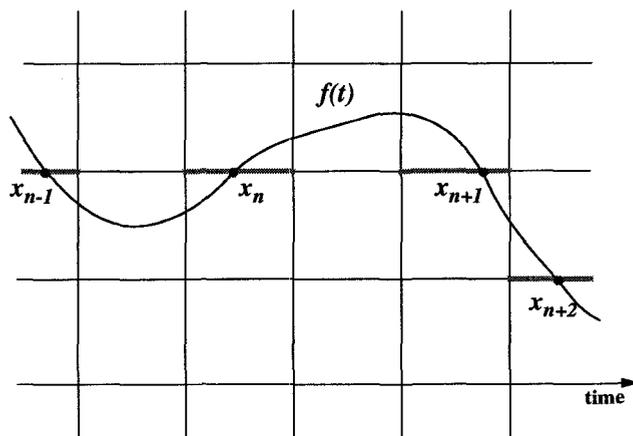


Fig. 4. Quantization threshold crossing types. A quantization threshold crossing can be immediately preceded by a crossing of the same quantization threshold, as illustrated by the crossings at points x_n and x_{n+1} . These are denoted as s -crossings, and each of them is preceded by a point where the consider signal assumes an extremum. The other type of quantization threshold crossings, d -crossings, are those which occur after a crossing of a different quantization threshold. The threshold crossing at x_{n+2} is of this type.

of a σ -bandlimited signal $f(t)$ on an interval T is the sum of numbers of these two types of crossings. The count of d -crossings, Q_d , depends on the slope of $f(t)$ as well as the quantization step q . The slope of $f(t)$ can be bounded as

$$|f'(t)| \leq \|f\| \sigma^{\frac{3}{2}}, \quad (8)$$

which gives

$$\frac{Q_d}{T} \leq \frac{\sigma^{\frac{3}{2}}}{q} \|f\|. \quad (9)$$

For the count of s -crossings, Q_s , we can investigate some average density. A sequence of points $\{\lambda_n\}$ on the real axis is said to have uniform density d if there exist two numbers $\delta > 0$ and $L < \infty$ such that

$$|\lambda_n - \frac{n}{d}| \leq L, \quad n = 0, \pm 1, \pm 2, \dots \quad (10)$$

$$|\lambda_n - \lambda_m| \geq \delta \quad n \neq m.$$

TABLE I

Error-rate characteristics of oversampled A/D conversion, as the sampling interval tends to zero, for the four different combinations of reconstruction and encoding. The quantization error, $e(t)$, is expressed as a function of the bit-rate, B .

		Linear Reconstruction
PCM		$E(e(t)^2) = O\left(\frac{1}{B}\right)$
Efficient Encoding		$E(e(t)^2) = O\left(2^{-\frac{T}{\sigma_m} B}\right)$
		Consistent Reconstruction
PCM		$\ e(t)\ ^2 = O\left(\frac{1}{B^2}\right)$
Efficient Encoding		$\ e(t)\ ^2 = O\left(2^{-2\frac{T}{\sigma_m} B}\right)$

Each of the s -crossings of $f(t)$ is preceded by a point where $f(t)$ assumes a local extremum. If s -crossings constitute a sequence of a uniform density d , then a subset of zeros of $f'(t)$ also constitutes a sequence of uniform density d . According to results on nonharmonic Fourier expansions by Duffin and Schaeffer [6], $d \leq \sigma/\pi$ except in a degenerate case when $f'(t)$ is identically equal to zero. If we assume that the sequence of quantization threshold crossings of $f(t)$ is a realization of an ergodic process, then as the interval T grows

$$\frac{Q_s}{T} \rightarrow c \leq \frac{\sigma}{\pi}. \quad (11)$$

Hence the error rate characteristics are given by

$$e(t)^2 = O\left(2^{-\frac{2}{\alpha_1 + \alpha_2} B}\right),$$

where $\alpha_1 \leq \|f\| \sigma^{\frac{3}{2}}/q$ and $\alpha_2 \leq \sigma/\pi$. Recall that for A/D conversion with a fixed sampling frequency f_s , $f_s \geq \sigma/\pi$, and $q \rightarrow 0$, the error-rate characteristics have the form

$$E(e(t)^2) = O\left(2^{-\frac{2}{\alpha} B}\right),$$

where $\alpha = f_s$.

It is interesting to find the error-rate characteristics of oversampled A/D conversion for the four combinations of reconstruction and encoding, i.e. linear versus consistent reconstruction, and PCM versus quantization threshold crossings encoding. These are given in Table 1. The results in Table 1 demonstrate the importance of appropriate lossless encoding of the digital representation. We can see that with consistent reconstruction and quantization threshold crossings encoding, error-rate characteristics of oversampled A/D conversion are significantly improved, and that for attaining an exponential decay efficient encoding is crucial.

V. CONCLUSION

This paper was concerned with error-rate characteristics of simple oversampled A/D conversion. An efficient scheme for lossless encoding of digital sequences produced in oversampled A/D conversion was introduced. It was shown that the error-rate characteristics of oversampled A/D conversion can be remarkably improved if consistent reconstruction and efficient coding are used, giving an error which behaves as $\|e(t)\|^2 = O(\exp(-2\beta B))$ rather than $E(e(t))^2 = O(1/B)$, which is the classical result.

REFERENCES

- [1] W. R. Bennett, "Spectra of Quantized Signals", Bell System Technical Journal, Vol.27, July 1948, pp.446-472.
- [2] N. T. Thao and M. Vetterli, "Reduction of the MSE in R-times oversampled A/D conversion from $O(1/R)$ to $O(1/R^2)$ ", IEEE Trans. on SP, Vol. 42, No. 1, pp. 200-203, January 1994.
- [3] Z. Cvetković and M. Vetterli, "Error Analysis in Oversampled A/D Conversion and Quantization of Weyl-Heisenberg Frame Expansions", submitted to IEEE Trans. on Information Theory, May 1995.
- [4] R. M. Young, "An Introduction to Nonharmonic Fourier Series", Academic Press, New York, 1980
- [5] S. Hein and A. Zakhor, "Reconstruction of Oversampled Band-Limited Signals from $\Sigma\Delta$ Encoded Binary Sequences", IEEE Trans. on Signal Processing, Vol.42, No.4, pp.799-811, April 1994.
- [6] R. J. Duffin and A. C. Schaeffer, "A Class of Nonharmonic Fourier Series", Trans. Amer. Math. Soc., Vol.72, March 1952, pp.341-366.