LOWER BOUNDS ON THE MSE IN N^{TH} ORDER MULTI-BIT MULTI-LOOP ΣΔ MODULATION WITH DC INPUTS

Nguyen T. Thao and Martin Vetterli *

Department of Electrical Engineering
and Center for Telecommunications Research
Columbia University, New York, NY 10027-6699

ABSTRACT

Recent research on non-linear decoding in ΣΔ modulation has shown that the classical linear decoding scheme is not optimal and that the MSE of reconstruction can be improved by 3 dB per octave of oversampling with a decoding scheme based on consistent estimation. In the example of multi-loop ΣΔ modulation, we present an analysis of the intrinsic behavior of an encoder, based on signal partitioning, thus providing a basic tool for the theoretical analysis of optimal decoding MSE. We show the application of this analysis tool to the case of MSE lower bounds for constant input signals in multi-loop ΣΔ modulation.

1. Introduction

The most successful technique currently used in oversampled A/D conversion(ADC) is ΣΔ modulation [1]. The samples of the bandlimited input signal are no longer purely quantized in amplitude, but processed through a complete encoding including integration, quantization and feedback. Figure 1 shows the double-loop encoding configuration of

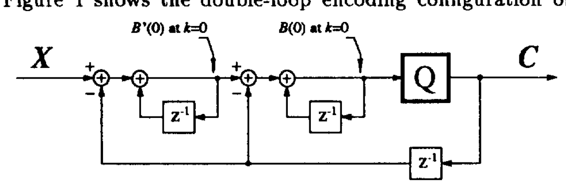

Figure 1: Block diagram of the double-loop ΣΔ modulator.

ΣΔ modulation. However, the reconstruction of the original input signal from its encoded version is classically performed like in the general case of oversampled ADC, that is, by using a lowpass filter. This is the decoding part of ΣΔ modulation and is basically a linear processing. The performance of this reconstruction is limited by the amount of in-band error existing in the encoded signal. This amount is well predicted using the classical white noise model of the quantization error signal [1]. In the multi-loop case, it was shown in [2] that the remaining error power decreases with the oversampling ratio R in $\mathcal{O}(R^{-(2n+1)})$, where n is the order of the modulator.

Questions were recently raised whether, given an encoding configuration, the linear decoding performs the best possible reconstruction of the input signal. The question of optimal reconstruction was first studied in [3, 4] in the case of constant input signals with the single-loop and the double-loop configurations. By studying the intrinsic behavior of the encoder, the optimal decoding of constant inputs was described and its non-linearity was demonstrated. It was proved in [4] that the mean square error (MSE) of the optimal decoding has lower bounds of the order $\mathcal{O}(R^{-3})$ and $\mathcal{O}(R^{-5})$ in single-loop and double-loop ΣΔ modulations respectively. This shows that, in the case of constant inputs, the optimal decoding MSE yields the same asymptotic dependence on R as the linear decoding MSE.

Non-linear decoding in the general context of oversampled ADC with time-varying bandlimited signals was first studied in [5]. The decoding criterion was to find, from the given encoded signal, a consistent estimate, that is, a bandlimited signal which necessarily produces the same encoded signal. An algorithm for consistent decoding, based on the principle of alternating projections, was proposed and tested [5]. When input signals are sinusoids, it was found that the achieved consistent estimates have an MSE dependence of the order $\mathcal{O}(R^{-4})$ and $\mathcal{O}(R^{-6})$ for the single-loop and the double-loop ΣΔ modulators, instead of $\mathcal{O}(R^{-3})$ and $\mathcal{O}(R^{-5})$ in linear decoding. This represented an asymptotic improvement of 3 dB per octave of oversampling over linear decoding for the two types of encoder. Using the same consistent decoding scheme, the asymptotic behavior of the order $\mathcal{O}(R^{-(2n+2)})$ instead of $\mathcal{O}(R^{-(2n+1)})$ in linear decoding, was obtained in [6, 7] with n^{th} order multi-bit multi-loop ΣΔ modulators, and n^{th} order multi-stage ΣΔ modulators, and more complex bandlimited input signals. Thus, contrary to the usual case in the classical analysis of oversampled ADC, the behavior of constant inputs is not representative for the general case.

To obtain general results on the theoretical limits of signal reconstruction in ΣΔ modulation, it is therefore necessary to deal with time-varying input signals. In this paper, we present a particular analysis of ΣΔ modulators in their multi-loop configuration, which points out a method to derive lower bounds on optimal decoding MSE with more complex signals than constant inputs. As a generalization of [3, 4], we show in Section 2 that optimal decoding is defined from the description of an encoder as inducing a partition of the space of input signals, whether they are constant, or more generally, bandlimited. For the multi-loop configuration of ΣΔ modulation, we show in Section 3 that this partition has very particular properties, independent of the nature of the input signals. As an application, we show in Section 4 how these properties can be used to derive MSE lower bounds when input signals are constant. In this application, we find a lower bound of the order $\mathcal{O}(R^{-(2n+2)})$ for multi-loop ΣΔ modulators of any order n, regardless of the number of bits used for the quantizer.

*Work supported in part by the National Science Foundation under grants ECD-88-11111.

0-7803-1254-6/93$03.00

2. Partitioning approach and optimal decoding

It was shown in [3, 4] that in the context of constant input signals, a $\Sigma\Delta$ modulator defines a one-to-one correspondence between intervals of dc input values and the possible encoded signals. Optimal decoding is therefore achieved by taking the center of the interval corresponding to the given encoded signal and its performance is an intrinsic characteristic of the encoder.

To generalize this approach to a broader class of signals, it is important to see that the encoder has an intrinsic behavior which is independent of the type of input signal. Based on the deterministic definition of quantization, it is shown in [6, 7] that the whole encoder works as a mapping, transforming a discrete-time input signal[1] X of the space $\mathbf{R}^N$ into an encoded signal C which is another signal of $\mathbf{R}^N$ (see Figure 1). This mapping naturally induces a partition of $\mathbf{R}^N$ where each cell $\mathcal{C}(C)$ corresponds to a possible encoded signal C and comprises all input signals X mapped into C. For illustration, we show in Figure 2 the partition induced by a two-bit double-loop $\Sigma\Delta$ modulator in the case where signals are limited to two samples ($N = 2$). On various types of encoders (simple, predictive, noise-shaping encoders), it was proved in [5, 6, 7] that the cells $\mathcal{C}(C)$ are necessarily convex. Thus, once the encoder is characterized, its intrinsic behavior in the context of oversampled ADC can be studied. In this context, input signals are confined to belong to a certain subspace $\mathcal{V}$ of bandlimited signals. By taking the restriction of the signal partition to the subspace $\mathcal{V}$, we naturally conclude that the encoder induces a partition on $\mathcal{V}$ composed of cells of the type $\mathcal{C}(C)\cap\mathcal{V}$ which are convex. If the input signals are constant, $\mathcal{V}$ is simply a one dimensional subspace of $\mathbf{R}^N$, and the cells $\mathcal{C}(C)\cap\mathcal{V}$ are necessarily segments of $\mathcal{V}$ (see Figure 2). This gives back the description presented in [3, 4] as a particular case. However, this description is valid for input spaces $\mathcal{V}$ of any dimension, and optimal decoding consists in general in taking the centroid of the cell $\mathcal{C}(C)\cap\mathcal{V}$, given the encoded signal C. Figure 3 shows the partition induced by a single-loop $\Sigma\Delta$ modulator, in the case where $\mathcal{V}$ is the two dimensional space of zero-phase sinusoids of fixed period and arbitrary dc component. For the case of constant inputs, the dc value intervals introduced in [3, 4] can be seen in Figure 3 by taking the intersection of the two dimensional partition with the dc component axis (represented by the dotted line).

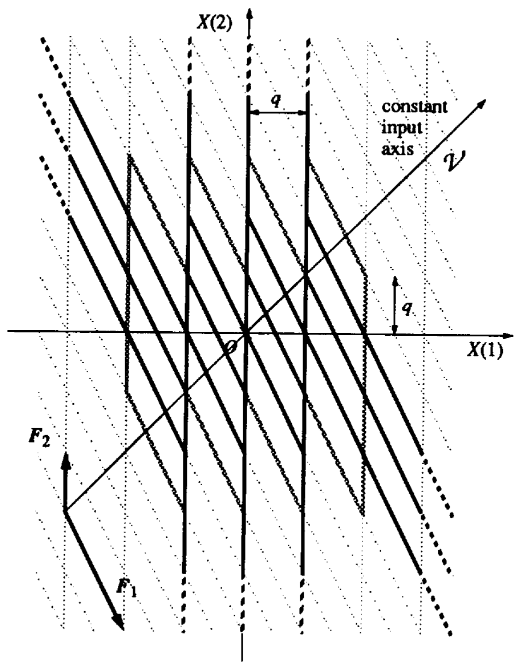

Figure 2: Partition induced in $\mathbf{R}^2$ by a two-bit double-loop $\Sigma\Delta$ modulator. The shaded lines represent the non-overload region and the dotted lines represent the case of infinite quantizer.

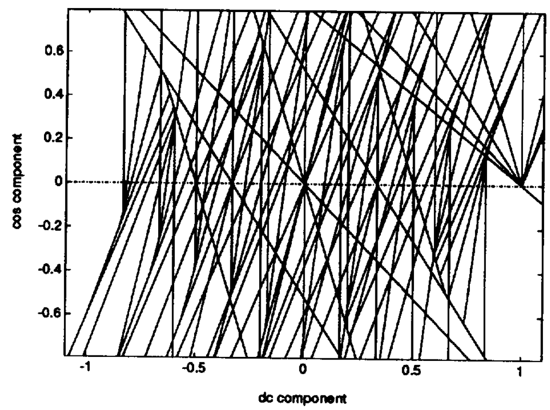

Figure 3: Partition induced on $\mathcal{V}$ by a single-loop $\Sigma\Delta$ modulator, where $\mathcal{V}$ is the two dimensional space of zero-phase sinusoids of fixed frequency and arbitrary dc component. The number of samples per period is 12.

Assuming that the input signals are constant with dc values uniformly distributed in an interval of length d, it was shown in [3] that a lower bound on the optimal decoding MSE can be obtained from the total number M of cells dividing the input interval. It was shown that the optimal MSE is necessarily greater than that obtained if the partition of the interval was uniform. This leads to [3]:

$$MSE_{opt} \geq \frac{d^2}{12M^2}, \tag{1}$$

The generalization of this technique to input signal spaces of higher dimension would be to count the number M per unit volume in $\mathcal{V}$, and take as a lower bound, the smallest MSE achievable by a partition, given the density M. For example, in two dimensions, the minimal MSE is achieved by hexagonal cells. Results for partitions of higher dimensions can be found in [8]. In order to derive the density of cells, it is important to first study the properties of the partition induced by the encoder.

3. Partition induced by a multi-loop $\Sigma\Delta$ modulator

In this section, we show that the partition induced on $\mathbf{R}^N$ by a multi-loop $\Sigma\Delta$ modulator has some interesting lattice properties. To analyze this partition, it is useful to work on the equivalent block diagram of a noise-shaping encoder shown in Figure 4 which was introduced in [2]. We show

[1] We assume that input signals after sampling are N-point sequences. The k^{th} sample of an input signal X is denoted by $X(k)$.

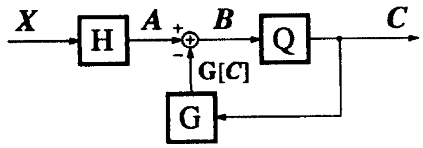

Figure 4: Equivalent block diagram of noise-shaping encoder [2].

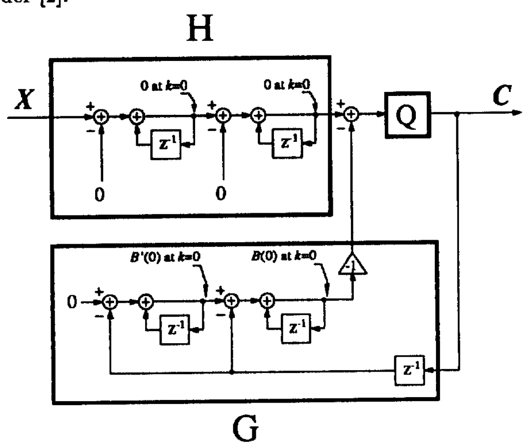

Figure 5: Equivalent block diagram of a double-loop $\Sigma\Delta$ modulator.

in Figure 5 how a double-loop $\Sigma\Delta$ modulator can be transformed to the equivalent diagram. In the general n-loop case, we see that **H** is an n^{th} order integrator and **G** is some deterministic feedback function. The procedure is the following. To start with, let us consider every possible encoded signal C, and derive the partition induced by the quantizer on the space of signals B (notations of Figure 4). Figure 6(a) shows the resulting partition in the case of a two-bit double-loop $\Sigma\Delta$ modulator using a uniform quantizer of step size q. If the quantizer had an infinite number of levels, the cell vertices would form a cubic lattice generated by the basis of signals $(E_j)_{1\leq j\leq N}$ defined by $E_j(k) = 0$ for all $k \neq j$, and $E_j(j) = q$. The lattice is represented by dotted lines in Figure 6(a). When the quantizer is finite, the induced partition is derived from the lattice by merging certain cells (those located outside the non-overload region represented by a shaded line). We will say that the partition has a lattice derived structure.

The second step is to derive the partition induced on the space of signals A (see Figure 4). The procedure is the following: for each encoded signal C, find the corresponding cell induced by the quantizer and translate it by the vector $\mathbf{G}[C]$. For the two-bit double-loop configuration, Figure 6(b) shows that the resulting partition has a lattice derived structure, with the same basis $(E_j)_{1\leq j\leq N}$. The reason is that the samples of the feedback signal $\mathbf{G}[C]$ are necessarily multiples of the lattice period q, since the operator **G** is limited to pure summations [7].

As a final step, the partition induced by the complete encoder is obtained by transforming the partition of signals A through the operator $\mathbf{H}^{-1}$, which is an n^{th} order differentiator. Since $\mathbf{H}^{-1}$ is linear, the lattice structure of the previously obtained partition, is transformed into another lattice structure (not necessarily cubic). The two-bit

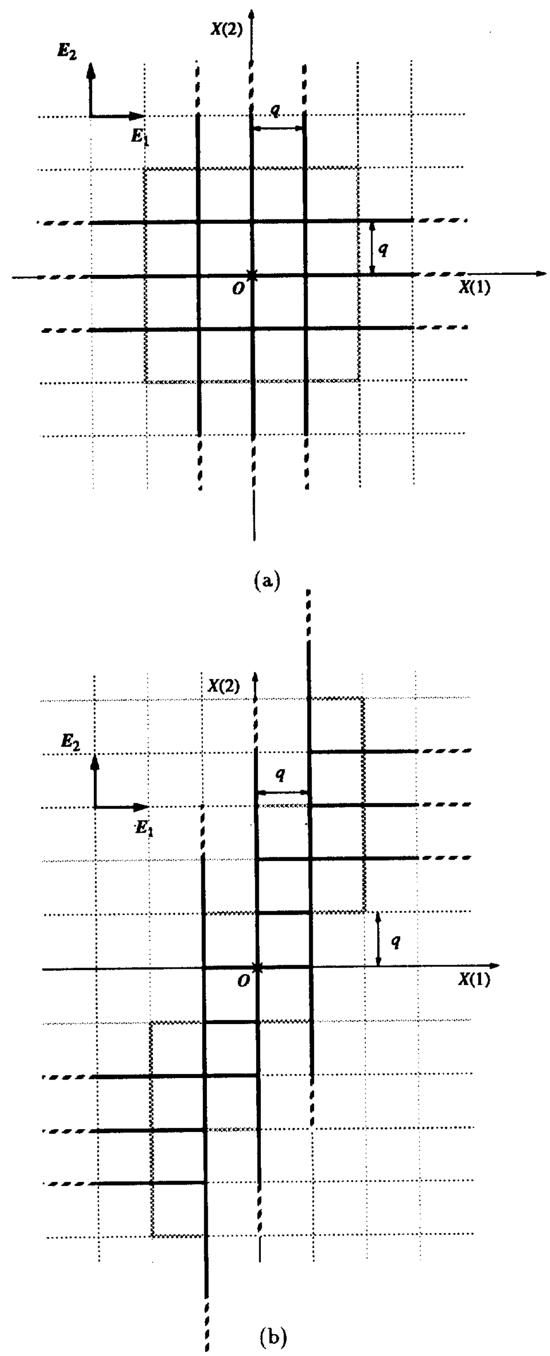

Figure 6: Intermediate partitions induced by a two-bit double-loop $\Sigma\Delta$ modulator (same line conventions as in Figure 2. (a) Partition induced at the node B of the equivalent block diagram (notations of Figure 4). (b) Partition induced at the node A of the equivalent block diagram.

double-loop case is shown in Figure 2. Therefore, the partition has a lattice derived structure, with basis $(F_j)_{1\leq j\leq N}$ where $F_j = \mathbf{H}^{-1}[E_j]$ for all $j = 1, ..., N$.

4. Application to the constant input case

As an example of application, we show how the lattice properties of the global partition can be used to derive a lower bound on the optimal decoding MSE in the case of constant inputs. Let us assume that input signals are uniformly distributed in a segment $[X_{min}, X_{max}]$ of $\mathcal{V}$ of length d. The inequality (1) of Section 2 shows that an MSE lower bound can be obtained by finding an upper bound on the number M of partition cells dividing the segment $[X_{min}, X_{max}]$ [3]. Let us assume for a while that the quantizer of the multi-loop $\Sigma\Delta$ modulator is infinite. We saw that, in this case, the partition induced by the encoder is a lattice generated by the basis $(F_j)_{1\leq j\leq N}$. In the two-bit double-loop case, Figure 7 shows the relative position of the segment $[X_{min}, X_{max}]$ with respect to the lattice. To find an upper

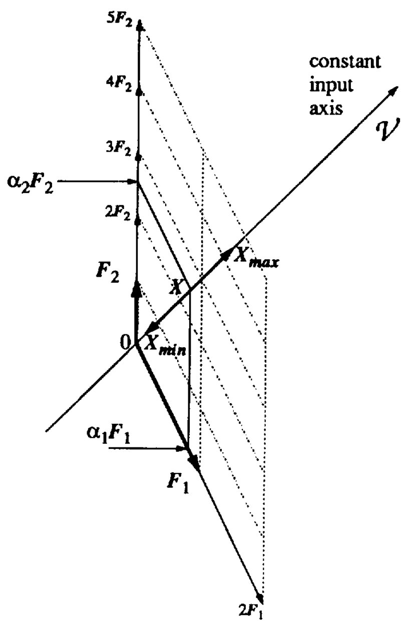

Figure 7: Relative position of the space $\mathcal{V}$ of constant input signal with the lattice structure of the partition of Fig. 2.

bound on M, it is first necessary to calculate this relative position. This amounts to calculating the decomposition of the spanning signal X_0 on the basis $(F_j)_{1\leq j\leq N}$:

$$X_0 = \sum_{j=1}^{N} \alpha_j^0 \cdot F_j.$$

Applying the linear operator $\mathbf{H}$ on this decomposition, we find $\mathbf{H}[X_0] = \sum_{j=1}^{N} \alpha_j^0 \cdot \mathbf{H}[F_j] = \sum_{j=1}^{N} \alpha_j^0 \cdot E_j$. According to the definition of $(E_j)_{1\leq j\leq N}$, E_j is simply the unit impulse located at the time index $k = j$. Therefore, α_j^0 is necessarily the value of the sequence $\mathbf{H}[X_0]$ at the time index $k = j$. But $\mathbf{H}[X_0]$ is the n^{th} order discrete-time integration of the constant signal X_0 of dc value 1. As a result, it can be easily shown that

$$\alpha_j^0 = \tfrac{1}{n!} j(j-1)\cdots(j+n-1) = \mathcal{O}(j^n).$$

Let us write $X_{min} = \sum_{j=1}^{N} \alpha_j^{min} \cdot F_j$ and $X_{max} = \sum_{j=1}^{N} \alpha_j^{max} \cdot F_j$, and let $X = \sum_{j=1}^{N} \alpha_j \cdot F_j$ be a signal of the segment $[X_{min}, X_{max}]$. Suppose that we move the signal X from X_{min} to X_{max}. From Figure 7 it is easy to see that whenever X enters a new cell of the lattice, there is at least one of the coefficients α_j which has its integer part increased by 1. By scanning the whole segment, one can derive that the total number M of encountered cells is upper bounded by

$$M \leq \sum_{j=1}^{N} \left(\lfloor\alpha_j^{max}\rfloor - \lfloor\alpha_j^{min}\rfloor + 1\right),$$

where $\lfloor x \rfloor$ designates the integer part of x. This was derived with the assumption that the quantizer is infinite. However, this inequality still holds when the quantizer is finite, since, by the merging of certain cells, M can only decrease. It is easy to see that $\lfloor\alpha_j^{max}\rfloor - \lfloor\alpha_j^{min}\rfloor \leq \alpha_j^{max} - \alpha_j^{min} + 1$. Moreover, because the segment $[X_{min}, X_{max}]$ is of length d, $X_{max} - X_{min} = d \cdot X$ which implies that $\alpha_j^{max} - \alpha_j^{min} = d \cdot \alpha_j^0$ for $j = 1, ..., N$. Then, we obtain

$$M \leq \sum_{j=1}^{N} d \cdot \alpha_j^0 + 2N = d \sum_{j=1}^{N} \mathcal{O}(j^n) + 2N = \mathcal{O}(N^{n+1}).$$

From (1), this finally implies that

$$MSE_{opt} \geq \mathcal{O}(N^{-(2n+2)})) = \mathcal{O}(R^{-(2n+2)}).$$

Thus, we have shown that the dependence of the optimal reconstruction MSE with R is lower bounded by $\mathcal{O}(R^{-(2n+2)})$, for any order n, and regardless of the number of bits of the quantizer. Note that in the first and second order case, stronger lower bounds were found in [4].

5. Conclusion and future work

In $\Sigma\Delta$ modulation, the optimal decoding is defined from the intrinsic behavior of the encoder which is to induce a partition of the space of discrete-time signals. We show, in the multi-loop configuration, that this partition has some particular lattice properties which constitute potential tools for the analysis of MSE lower bounds in the general case of time-varying bandlimited signals. As a first demonstration, we show their immediate application in the case of constant inputs. The present partition analysis will be used to study MSE lower bounds with sinusoidal inputs in a future work.

References

[1] J. Candy and G.C.Temes, eds., *Oversampling delta-sigma data converters. Theory, design and simulation.* IEEE Press, 1992.

[2] S.K.Tewksbury and R.W.Hallock, "Oversampled, linear predictive and noise shaping coders of order $N > 1$," *IEEE Trans. Circuits and Systems*, vol. CAS-25, pp. 436–447, July 1978.

[3] A.Zakhor and K.Ibraham, "Lower bounds on the MSE of the single loop sigma delta modulator," *Proc. 23th Asilomar Conf. in Signals, Systems and Computers*, pp. 849–853, Oct. 1989.

[4] S.Hein and A.Zakhor, "Lower bounds on the MSE of the single and double loop sigma delta modulators," *Proc. IEEE Int. Symp. Circ. and Systems*, pp. 1751–1755, May 1990.

[5] N.T.Thao and M.Vetterli, "Oversampled A/D conversion using alternate projections," *Conf. on Information Sciences and Systems, the Johns Hopkins University*, pp. 241–248, Mar. 1991.

[6] N.T.Thao and M.Vetterli, "Deterministic analysis of oversampled A/D conversion and decoding improvement based on consistent estimates," *IEEE Trans. on Signal Proc.* Submitted.

[7] T.T.Nguyen, "Deterministic analysis of oversampled A/D conversion and $\Sigma\Delta$ modulation, and decoding improvements using consistent estimates," *PhD. dissertation, Dept. of Elect. Eng., Columbia Univ.*, Feb. 1993.

[8] J.H.Conway and N.J.A.Sloane, *Sphere Packings, Lattices and Groups.* Springler-Verlag, 1988.