Time Average MSE Analysis for the First Order Sigma-Delta Modulator with the Inputs of Bandlimited Signals

Wen Chen
Department of Electrical and Computer Engineering
University of Alberta, Edmonton, AB, Canada, T6G 2V4
Email: wenchen@ece.ualberta.ca

Chintha Tellambura
Department of Electrical and Computer Engineering
University of Alberta, Edmonton, AB, Canada, T6G 2V4
Email: chintha@ece.ualberta.ca

***Abstract*— Based on experiments and numerical simulation, it has been widely believed that the time average mean square error in the first order sigma-delta modulator with input of bandlimited signals decays like $O(\lambda^{-3})$ as the sampling ratio λ goes to infinity. This conjecture remains as an open problem for many years. Combining tools from number theory, harmonic analysis, real analysis and complex analysis, this paper shows that the conjecture holds in some reasonable sense.**

I. Introduction

Converting an analog signal into a digital signal (A/D conversion) consists of two steps: *sampling* and *quantization*. According to sampling theorems [3], [4], [7], [16], the loss in the sampling step is reversible. However, the quantization step usually introduces an irreversible loss of information. The loss in the quantization step is called the *quantization error*. Recent research shows that the accuracy of A/D conversion can be improved by refining the resolution of sampling as well as quantization [3], [6]. Due to the expensive cost in building a high resolution quantizer, high accuracy of A/D conversion is usually achieved by refining the resolution of sampling. Among many schemes, the sigma-delta ($\Sigma\Delta$) modulator provides a promising architecture for high accuracy A/D conversion because it is robust against circuit imperfections and hence is amenable to LSI and VLSI imperfection [1], [2], [7].

In the literature, the input signal x to a $\Sigma\Delta$ modulator is modelled to be a bandlimited function taking values in the interval $[0,1]$ or $[-1,1]$ for the convenience of discussion [7], [9], [10], [11], [12]. For the same reason, in this paper, we consider the signal x to be a bandlimited function taking values in the interval $[1,2]$ instead. But these assumptions are essentially equivalent. In this paper, we consider the bandlimited signals in the space $L^\infty(\mathbb{R})$ so that some important signals, such as sinusoidal signals, can be included in our discussion. In other words, we define the signal class to be

$$\mathcal{B}_\Omega := \left\{ x \in L^\infty(\mathbb{R}) \ \middle|\ \hat{x} \text{ has support contained in } [-\Omega,\Omega] \right\},$$

where $\hat{x}$ denotes the Fourier transform of x which is defined by $\hat{x}(\xi) := \int_{-\infty}^{\infty} x(t)e^{-i\xi t}dt$ for $x \in L^1(\mathbb{R})$ and extended to the tempered distributions in the usual way [8]. Without loss of generality, in this paper we work with the space $\mathcal{B}_\pi$ and all the results in this paper can be easily extended to arbitrary bandwidth Ω by rescaling. For $\lambda > 1$, let φ be a function in the Schwartz class such that $\hat{\varphi} = 1$ on the interval $[-\pi,\pi]$ and $\hat{\varphi}$ has support contained in the interval $[-\lambda\pi,\lambda\pi]$. Then the following sampling formula [9], [10], [14] holds

$$x(t) = \frac{1}{\lambda}\sum_{n\in\mathbb{Z}} x\left(\frac{n}{\lambda}\right)\varphi\left(t-\frac{n}{\lambda}\right), \qquad x \in \mathcal{B}_\pi. \tag{I.1}$$

This φ is called an *admissible reconstruction filter*, the quantity λ in the above sampling formula is called the *sampling ratio*, and $\{x(n/\lambda)\}_{n\in\mathbb{Z}}$ are called the *samples* of the input signal x.

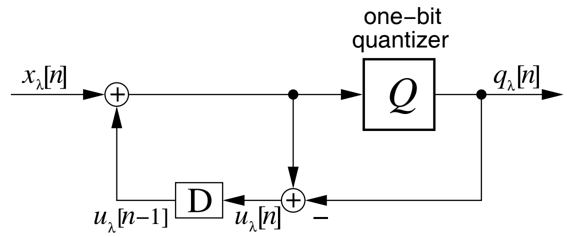

Fig. 1. A basic $\Sigma\Delta$ modulator using (I.6) and (I.7).

For $t \in \mathbb{R}$, we denote by $\lfloor t \rfloor$ the largest integer that is not greater than t and denote $\langle t \rangle := t - \lfloor t \rfloor \in [0,1)$. The $\Sigma\Delta$ modulator uses the samples $\{x(n/\lambda)\}_{n\in\mathbb{Z}}$ as inputs to generate the binary digital signal $\{q_\lambda(n)\}_{n\in\mathbb{Z}}$ as follows:

$$X_\lambda(n) := \sum_{m=1}^{n} x\left(\frac{m}{\lambda}\right), \tag{I.2}$$

$$Q_\lambda(n) := \lfloor X_\lambda(n) \rfloor, \tag{I.3}$$

$$q_\lambda(n) := Q_\lambda(n) - Q_\lambda(n-1), \tag{I.4}$$

where $X_\lambda(0) := 0$ and $X_\lambda(n) := -\sum_{m=n+1}^{0} x(m/\lambda)$ for $n < 0$. It is easy to see that $X_\lambda(n) - X_\lambda(n-1) = x(n/\lambda)$ for all $n \in \mathbb{Z}$. Since x takes value in the interval $[1,2]$, it is easy to check that $q_\lambda(n)$ takes the binary value 1 or 2. Note that the equations (I.2) and (I.4) correspond to "Σ" and "Δ", respectively; hence giving the name of the modulator. Since X_λ and Q_λ will accumulate into huge numbers as time elapses, neither can be calculated in a circuit. Thus one introduces the auxiliary variable $u_\lambda := X_\lambda - Q_\lambda = \langle X_\lambda \rangle$. Then u_λ satisfies the recursive relation:

$$u_\lambda(n) - u_\lambda(n-1) = x\left(\frac{n}{\lambda}\right) - q_\lambda(n). \tag{I.5}$$

Since $u_\lambda(n) \in [0,1)$, from (I.5) one has the relation $q_\lambda(n) = \lfloor x(n/\lambda) + u_\lambda(n-1) \rfloor$. Using the auxiliary variable u_λ, now we can translate the procedure in (I.2)–(I.4) into the following equivalent procedure:

$$u_\lambda(n) := u_\lambda(n-1) + x\left(\frac{n}{\lambda}\right) - q_\lambda(n) \qquad \text{with} \quad u_\lambda(0) := 0, \tag{I.6}$$

$$q_\lambda(n) := \begin{cases} 1, & \text{if} \quad x\left(\frac{n}{\lambda}\right) + u_\lambda(n-1) < 2; \\ 2, & \text{if} \quad x\left(\frac{n}{\lambda}\right) + u_\lambda(n-1) \geqslant 2. \end{cases} \tag{I.7}$$

The simplest $\Sigma\Delta$ modulator is shown in Figure 1 and consists of a discrete time integrator and a binary quantizer inside a single feedback loop [7], [12], [18].

Reconstructing the analog signal from the digital sequence $\{q_\lambda(n)\}_{n\in\mathbb{Z}}$ by the sampling formula in (I.1), one obtains the reconstructed signal

$$\tilde{x}(t) := \frac{1}{\lambda}\sum_{n\in\mathbb{Z}} q_\lambda(n)\varphi\left(t-\frac{n}{\lambda}\right).$$

Then the quantization error is defined to be the difference $e_\lambda(x,t) := x(t)-\tilde{x}(t)$. The basic estimate given in [7], [9] for the quantization error is

$$\|e_\lambda(x,\cdot)\|_\infty \leqslant \|\varphi'\|_1\lambda^{-1}.$$

In practice, however, one observes a much better decay behaviour than that of this basic estimate. In particular, through experiments and numerical simulation, it is commonly believed that

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}|e_\lambda(x,t)|^2dt \leqslant C\lambda^{-3} \tag{I.8}$$

with C independent of the input signal x.

However, it is turn out to be very hard to improve the basic estimate [7]. For very restricted classes of signals x, the conjecture in (I.8) has been proved in some sense. In particular, Gray [11] showed that if $x_a = a$ for $a \in [-1,1]$, then

$$\int_{-1}^{1}\left[\lim_{T\to\infty}\frac{1}{T}\int_{-T}^{T}|e_\lambda(x_a,t)|^2dt\right]da \leqslant C\lambda^{-3}.$$

Gray's result was later extended by Gray, Chou and Wong [12] to the case where the input signal x is sinusoid; that is, $x(t) = a\sin bt$, with $|b| < \pi$.

For general bandlimited signals, to our best knowledge, there was little progress on this conjecture until the recent interesting work of Güntürk [9], [10], who proved, by a combination of tools from number theory and harmonic analysis, that if $x \in \mathcal{B}_\pi$ takes values in the interval $[0,1]$, then for any small $\varepsilon > 0$ and all t satisfying $x'(t) \neq 0$, there is a constant $C_{\varepsilon,x'(t)} > 0$ such that

$$|e_\lambda(x,t)|^2 \leqslant C_{\varepsilon,x'(t)}\lambda^{-8/3+\varepsilon},$$

where the constant $C_{\varepsilon,x'(t)}$ depends on ε and the local value $x'(t)$.

Our work in this paper has been greatly motivated and inspired by the recent interesting work in Daubechies and DeVore [7] and Güntürk [9], [10]. Our objective in this paper is to try to prove the conjecture in (I.8). Using tools from number theory, harmonic analysis, real analysis and complex analysis, we shall prove the conjecture in the sense that the time average MSE decays like $O(\lambda^{-3+\epsilon})$ for any small $\epsilon > 0$ as the sampling ratio λ approaches infinity. Furthermore, if the signals are restricted to some special classes including the periodic signals, our result is very much close to the original conjecture in (I.8). See the next section for details of the main results in this paper.

II. Main Results

We now present the main results in this paper. Since the conjecture in (I.8) has been confirmed by Gray [11] and Güntürk [9], [10] for a constant input, in the following we only deal with a non-constant signal $x \in \mathcal{B}_\pi$.

Theorem 1: Suppose that $x \in \mathcal{B}_\pi$ takes values in the interval $[1,2]$. Then for any large $T > 0$ and any small $\varepsilon > 0$, there are two positive constants $\lambda_{x,T}$ and $C_{\varepsilon,x,T}$ such that

$$\frac{1}{2T}\int_{-T}^{T}|e_\lambda(x,t)|^2dt \leqslant C_{\varepsilon,x,T}\lambda^{-3+\varepsilon} \qquad \forall\,\lambda > \lambda_{x,T},$$

where $C_{\varepsilon,x,T}$ depends on ε, T and the property of the signal x.

Theorem 1 implies that, considering the input signal x on any finite time interval $[-T,T]$, the decay of the average square error behaves like $O(\lambda^{-3+\varepsilon})$ for any small $\varepsilon > 0$. Since ε can be arbitrarily small, Theorem 1 is close to the conjecture in (I.8) on a finite time interval. Since in experiments and numerical simulation one always deals with signals on a finite time interval, Theorem 1 explains well the observations made in experiments.

The constant $C_{\varepsilon,x,T}$ in Theorem 1 is determined by the behaviour of the signal x in the interval $(-T-\tau, T+\tau)$ for any fixed $\tau > 0$. Suppose that the increasing sequence $\{\alpha_v\}_{v\in\Lambda}$ are all the distinct zeros of x'' on $\mathbb{R}$, where Λ is a set of consecutive integers. Suppose also that $\{\alpha_v\}_{v=-M(T)}^{N(T)}$ are all the distinct zeros of x'' on $[-T,T]$. If there is a zero of x'' in the open interval $(-T-\tau,-T)$, we set $T_b := \alpha_{-M(T)-1}$ and $\sigma_b := -M(T)-1$, otherwise $T_b := -T$ and $\sigma_b := -M(T)$; If there is a zero of x'' in the open interval $(T, T+\tau)$, we set $T_B := \alpha_{N(T)+1}$ and $\sigma_B := N(T)+1$, otherwise $T_B := T$ and $\sigma_B := N(T)$. Let r_v denote the order of the zero α_v of the analytic function x'' for all $v \in \Lambda$, and $r_T := \sup_{\sigma_b\leqslant v\leqslant\sigma_B} r_v$. Let $m_T := \min_{\sigma_b\leqslant v\leqslant\sigma_B}|x^{(r_v+2)}(\alpha_v)|$. Then $m_T > 0$. For any positive small $d < \tau$ such that the d-neighbourhood of α_v defined by $V(\alpha_v,d) := (\alpha_v - d, \alpha_v + d)$ are disjoint to each other for $v = \sigma_b,\ldots,\sigma_B$, let

$$c_T(d) := \inf_{t\in[T_b-d/4,T_B+d/4]\setminus\cup_{v=\sigma_b}^{\sigma_B}V(\alpha_v,d/2)}|x''(t)|.$$

Then $c_T(d) > 0$. As a matter of fact, the constant $C_{\varepsilon,x,T}$ in Theorem 1 is determined by the quantities ε, r_T, m_T and $c_T(d)$. Naturally, if we put some conditions on the quantities r_T, m_T and $c_T(d)$, the result in Theorem 1 can be improved and we have the following theorem.

Theorem 2: Suppose that $x \in \mathcal{B}_\pi$ takes values in the interval $[1,2]$ and $\sup_{v\in\Lambda} r_v < \infty$. If there is a real number $d > 0$ such that $V(\alpha_v,d)$ are disjoint to each other for all $v \in \Lambda$,

$$\inf_{v\in\Lambda}\inf_{t\in V(\alpha_v,d)}|x^{(r_v+2)}(t)| > 0 \quad\text{and}\quad \inf_{t\in\mathbb{R}\setminus\cup_{v\in\Lambda}V(\alpha_v,d)}|x''(t)| > 0,$$

then for any small $\varepsilon > 0$, there exist two positive constants λ_x and $C_{\varepsilon,x}$ such that

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}|e_\lambda(x,t)|^2dt \leqslant C_{\varepsilon,x}\lambda^{-3+\varepsilon} \qquad \forall\,\lambda > \lambda_x,$$

where the constant $C_{\varepsilon,x}$ is independent of T and λ.

For any periodic signal, the quantities r_T, m_T and $c_T(d)$ are completely determined by its behaviour in one period. Therefore, r_T, m_T and $c_T(d)$ are independent of T. Hence we have the following result.

Corollary 1: Suppose that $x \in \mathcal{B}_\pi$ is periodic and takes values in the interval $[1,2]$. Then for any small $\varepsilon > 0$, there exist two positive constants λ_x and $C_{\varepsilon,x}$ such that

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}|e_\lambda(x,t)|^2dt \leqslant C_{\varepsilon,x}\lambda^{-3+\varepsilon} \qquad \forall\,\lambda > \lambda_x.$$

The sinusoidal signals considered by Gray *et al* [12] are obviously covered by Corollary 1. However, the constant $C_{\varepsilon,x}$ in Theorem 2 and Corollary 1 is not independent of the input x. We guess that such an absolute constant may depend on the statistical properties of the signals in $\mathcal{B}_\pi$. In Gray's estimate, he got the absolute constant since he took the integration over all possible signals $x_a = a\sin bt$ for $a \in [-1,1]$. It could be possible to define a measure in $\mathcal{B}_\pi$ and take integration of the average square error with respect to the measure to obtain an absolute constant.

III. Proof of Theorem 1

In this section, we will present the proof of Theorem 1. Since Theorem 2 and Corollary 1 are immediate consequences of our proof of Theorem 1, we will omit their proofs. In this section, we always use C to stand for a general absolute constant.

Before proceeding further, we need some auxiliary results. First, we review some important facts about bandlimited functions in [19].

- A bandlimited function $x \in \mathcal{B}_\Omega$ is the restriction to $\mathbb{R}$ of an entire function of exponential type Ω [19, page 85]; that is, the function defined by
$$x(z) = \frac{1}{2\pi} \int_{-\Omega}^{\Omega} e^{i\xi z} \hat{x}(\xi) d\xi, \quad z \in \mathbb{C},$$
is an entire function and satisfies $|x(z)| \leqslant Ae^{\Omega|z|}$ for all $z \in \mathbb{C}$ with some constant $A > 0$.
- Bernstein's inequality [19, page 88]: If $x \in \mathcal{B}_\Omega$, then $\|x^{(r)}\|_\infty \leqslant \Omega^r \|x\|_\infty$.

For a function x of complex variable, z_0 is called a zero of x if $x(z_0) = 0$. An interesting result about the zeros of an entire function of exponential type which plays an important role in this paper is as follows.

Proposition 1: (see [19, page 52]) If $x(z)$ is a nonzero entire function of exponential type Ω, then
$$\limsup_{T \to \infty} \frac{N(T)}{T} \leqslant \frac{2\Omega}{\ln 2},$$
where $N(T)$ is the number of zeros of x inside the disc with radius T and center at the origin.

The following fundamental estimate in number theory was used in Güntürk's work [9], [10] and will be useful in this paper.

Lemma 1: (Koksma-Erdös-Turán [13], [15]) Let f be a function of bounded variation on $[0, 1]$. For any finite sequence of points $u_1, \ldots, u_n$ in $[0, 1]$,
$$\begin{aligned} &\left| \frac{1}{n} \sum_{\ell=1}^{n} f(u_l) - \int_0^1 f(u) du \right| \\ \leqslant \quad & C V_0^1(f) \inf_{K \geqslant 1} \left(\frac{1}{K} + \sum_{k=1}^{K} \frac{1}{k} \left| \frac{1}{n} \sum_{\ell=1}^{n} e^{2\pi i k u_\ell} \right| \right), \end{aligned}$$
where $V_0^1(f)$ is the total variation of f on $[0, 1]$.

We also need an estimate for oscillatory integrals in harmonic analysis.

Lemma 2: (van der Corput [17, page 332]) Suppose that β is a real-valued function on the interval (a, b) such that for a positive integer r, $|\beta^{(r)}(t)| \geqslant \mu$ for all $t \in (a, b)$. Then
$$\left| \int_a^b e^{i\beta(t)} dt \right| \leqslant C_r \mu^{-1/r}, \tag{III.1}$$
holds when:

(i) $r \geqslant 2$ or

(ii) $r = 1$ and β' is monotonic in (a, b).

The constant C_r depends only on r.

The following result is an extension of an interesting result established by Güntürk [9], [10], which will be useful in this paper.

Lemma 3: (Güntürk [9], [10]) The sequence X_λ can be extended to an analytic function, which we shall denote by X_λ as well. Moreover, for any positive integer r,
$$X_\lambda^{(r)} = \frac{1}{\lambda^{r-1}} x^{(r-1)} \left(\frac{\cdot}{\lambda} \right) + R_{r,\lambda}$$
and
$$\|R_{r,\lambda}\|_\infty \leqslant \frac{(\pi/\lambda)^r e^{\pi/\lambda} \|x\|_\infty}{1 - \pi/\lambda}.$$

Now we present the proof of Theorem 1. For any $t \in \mathbb{R}$, define
$$N_\lambda(t) := \lfloor \lambda t \rfloor, \quad t_\lambda := \frac{N_\lambda(t)}{\lambda} \quad \text{and} \quad \delta_\lambda(t) := t - t_\lambda = \frac{\langle \lambda t \rangle}{\lambda}.$$
Let φ be a function in the Schwartz class such that $\hat{\varphi} = 1$ on $[-\pi, \pi]$ and has support contained in $[-2\pi, 2\pi]$. Therefore, there are constants $C_{r,N} > 0$ such that for all $N \geqslant 0$ and $r \geqslant 0$,
$$|\varphi^{(r)}(t)| \leqslant \frac{C_{r,N}}{(1 + |t|)^N} \qquad \forall\, t \in \mathbb{R}.$$
For a small $\varepsilon > 0$, set $\Omega_\lambda = \lambda^{\varepsilon/8}$ and define φ_λ by
$$\varphi_\lambda(t) := \Omega_\lambda \varphi(\Omega_\lambda t).$$
Then $\hat{\varphi}_\lambda(\xi) = \hat{\varphi}(\xi/\Omega_\lambda)$. Since $\hat{\varphi}_\lambda = 1$ on $[-\Omega_\lambda \pi, \Omega_\lambda \pi] \supset [-\pi, \pi]$ and $\hat{\varphi}_\lambda$ has support contained in $[-2\Omega_\lambda \pi, 2\Omega_\lambda \pi] \subset [-\lambda\pi, \lambda\pi]$ for $\lambda > 4$, $\{\varphi_\lambda\}_\lambda$ is a family of admissible reconstruction filters. Define the difference operator Δ_s by $\Delta_s \varphi := \varphi - \varphi(\cdot - s)$ and denote $\Delta := \Delta_1$. From the sampling formula in (I.1) and the definition of quantization error, by $t = t_\lambda + \delta_\lambda(t)$ and $\Delta u_\lambda(n) = x(n/\lambda) - q_\lambda(n)$, we have
$$\begin{aligned} e_\lambda(x, t) &= \frac{1}{\lambda} \sum_{n \in \mathbb{Z}} (x_\lambda(n/\lambda) - q_\lambda(n)) \varphi_\lambda \left(t - \frac{n}{\lambda} \right) \\ &= \frac{1}{\lambda} \sum_{n \in \mathbb{Z}} \Delta u_\lambda(n) \varphi_\lambda \left(t - \frac{n}{\lambda} \right) \\ &= \frac{1}{\lambda} \sum_{n \in \mathbb{Z}} u_\lambda(n) \Delta_{\frac{1}{\lambda}} \varphi_\lambda \left(t - \frac{n}{\lambda} \right) \\ &= \frac{1}{\lambda} \sum_{n \in \mathbb{Z}} u_\lambda(n + N_\lambda(t)) \Delta_{\frac{1}{\lambda}} \varphi_\lambda \left(\delta_\lambda(t) - \frac{n}{\lambda} \right). \end{aligned}$$
Since $u_\lambda \in [0, 1)$, we deduce that
$$\begin{aligned} h_\lambda(t) \quad := \quad & \frac{1}{\lambda} \sum_{n \in \mathbb{Z}} u_\lambda(n + N_\lambda(t)) \\ & \left[\Delta_{\frac{1}{\lambda}} \varphi_\lambda \left(\delta_\lambda(t) - \frac{n}{\lambda} \right) - \Delta_{\frac{1}{\lambda}} \varphi_\lambda \left(-\frac{n}{\lambda} \right) \right] \end{aligned}$$
satisfies
$$\begin{aligned} |h_\lambda(t)| &\leqslant \frac{1}{\lambda} \sum_{n \in \mathbb{Z}} \left| \Delta_{\frac{1}{\lambda}} \varphi_\lambda \left(\delta_\lambda(t) - \frac{n}{\lambda} \right) - \Delta_{\frac{1}{\lambda}} \varphi_\lambda \left(-\frac{n}{\lambda} \right) \right| \\ &\leqslant \frac{1}{\lambda} \sum_{n \in \mathbb{Z}} \int_{-n/\lambda}^{-n/\lambda + \delta_\lambda(t)} \left| \Delta_{\frac{1}{\lambda}} [\varphi_\lambda(s)]' \right| ds \\ &\leqslant \frac{1}{\lambda} \int_{-\infty}^{\infty} ds \int_{s-1/\lambda}^{s} |[\varphi_\lambda]''(\tau)| d\tau \\ &= \frac{1}{\lambda^2} \|[\varphi_\lambda]''\|_1 \\ &= \lambda^{-2+\varepsilon/4} \|\varphi''\|_1. \end{aligned}$$
Noting that $N_\lambda(t) = \lambda t_\lambda$, we rewrite the quantization error $e_\lambda(x, \cdot)$ as
$$\begin{aligned} e_\lambda(x, t) &= \frac{1}{\lambda} \sum_{n \in \mathbb{Z}} u_\lambda(n + N_\lambda(t)) \Delta_{\frac{1}{\lambda}} \varphi_\lambda \left(-\frac{n}{\lambda} \right) + h_\lambda(t) \\ &= \frac{1}{\lambda} \sum_{n \in \mathbb{Z}} \left[u_\lambda(n + \lambda t_\lambda) - \frac{1}{2} \right] \Delta_{\frac{1}{\lambda}} \varphi_\lambda \left(-\frac{n}{\lambda} \right) + h_\lambda(t). \end{aligned}$$
Since X_λ has been extended to an analytic function in Lemma 3, u_λ can be also extended to a function defined on the real line $\mathbb{R}$ by $u_\lambda(t) := \langle X_\lambda(t) \rangle, t \in \mathbb{R}$. Define an auxiliary symbol $U_\lambda(n, t)$ by
$$U_\lambda(n, t) := \begin{cases} \sum_{\ell=1}^{n} [u_\lambda(\lambda t + \ell) - 1/2] & n > 0; \\ 0, & n = 0; \\ -\sum_{\ell=n+1}^{0} [u_\lambda(\lambda t + \ell) - 1/2], & n < 0. \end{cases}$$

Then $u_\lambda(\lambda t+n)-1/2=\Delta U_\lambda(n,t)$. Therefore,

$$\begin{aligned} e_\lambda(x,t) &= \frac{1}{\lambda}\sum_{n\in\mathbb{Z}} \Delta U_\lambda(n,t_\lambda)\Delta_{\frac{1}{\lambda}}\varphi_\lambda\left(-\frac{n}{\lambda}\right)+h_\lambda(t) \\ &= \frac{1}{\lambda}\sum_{n\in\mathbb{Z}} U_\lambda(n,t_\lambda)\Delta^2_{\frac{1}{\lambda}}\varphi_\lambda\left(-\frac{n}{\lambda}\right)+h_\lambda(t) \\ &= \frac{1}{\lambda}\sum_{n>0}+\frac{1}{\lambda}\sum_{n=0}+\frac{1}{\lambda}\sum_{n<0} U_\lambda(n,t_\lambda)\Delta^2_{\frac{1}{\lambda}}\varphi_\lambda\left(-\frac{n}{\lambda}\right)+h_\lambda(t) \\ &=: e^+_\lambda(x,t)+0+e^-_\lambda(x,t)+h_\lambda(t). \end{aligned}$$

In order to estimate $e^+_\lambda(x,t)$, it suffices to estimate $U_\lambda(n,t_\lambda)$. Since $|U_\lambda(n,t_\lambda)|\leqslant|U_\lambda(n,t)|+|U_\lambda(n,t)-U_\lambda(n,t_\lambda)|$, we only need to estimate $|U_\lambda(n,t)|$ and $|U_\lambda(n,t)-U_\lambda(n,t_\lambda)|$, respectively. Since each step involves complicated tools from number thory, harmonic analysis, complex analysis and real analysis, we only list the major results in each steps while omit the detailed proofs.

1) Step 1: Estimate $|U_\lambda(n,t)|$. We obtain

$$\frac{1}{2T}\int_{-T}^{T}|U_\lambda(n,t)|^2\,dt\leqslant CC_{r_T}\left(r_T^2+b_T^{-1}\right)\lambda\ln^2\lambda, \tag{III.2}$$

for all $\lambda>16\pi^{r_T+4}/b_T$ and $n\leqslant\rho_T\lambda$.

2) Step 2: Estimate $|U_\lambda(n,t)-U_\lambda(n,t_\lambda)|$. We obtain

$$\frac{1}{2T}\int_{-T}^{T}|U_\lambda(n,t)-U_\lambda(n,t_\lambda)|^2dt\leqslant CC_{r_T}r_T^2\lambda\ln^2\lambda+CC_{r_T}b_T^{-1}\lambda.$$

3) Step 3: Final estimate. Let $\lambda_{x,T}=16\pi^{r_T+4}/b_T$. Combining Steps 1 and 2, by the triangle inequality we obtain that for all $\lambda>\lambda_{x,T}$ and $n\leqslant\rho_T\lambda$,

$$\frac{1}{2T}\int_{-T}^{T}|U_\lambda(n,t_\lambda)|^2dt\leqslant C_{x,T}\lambda\ln^2\lambda, \tag{III.3}$$

where $C_{x,T}:=CC_{r_T}(r_T^2+b_T^{-1})$. Using the admissible filter, Finally, we have

$$\frac{1}{2T}\int_{-T}^{T}|e^+_\lambda(x,t)|^2dt\leqslant C_{\varepsilon,x,T}\lambda^{-3+\varepsilon}.$$

Similarly, the same estimate for $e^-(x,\cdot)$ can be attained by reflecting the negative indices into the positive indices, which completes the proof of Theorem 1.

IV. Conclusion

In this paper, we regoriously analyze the time average mean square error of a sigma-delta system with the inputs of band-limited signals. Combining tools from number theory, harmonic analysis, real analysis and complex analysis, we establish the estimate $O(\lambda^{-3+\epsilon})$ for any small $\epsilon>0$ as the sampling ratio λ approaches infinity, which is very close to the long standing conjecture $O(\lambda^{-3})$.

Acknowledgment

The authors would like to thank Professor Nguyen T. Thao for his inspiration on the work done in this paper and Professor Bin Han and Rong-Qing Jia for their numerous support.

References

[1] J. C. Candy, A use of limit cycle oscillation to obtain robust analog-to-digital converters, *IEEE Trans. Communications,* **22** (1974), no. 3, 298–305.

[2] J. C. Candy and Y. Benjamin, The structure of quantization noise from sigma-delta modulation, *IEEE Trans. Communications,* **29** (1981), no. 9, 1316–1323.

[3] W. Chen, B. Han and R. -Q Jia, On simple oversampled A/D conversion in shift invariant spaces, to appear in *IEEE Trans. Information Theory,* **51** (2005), no. 2.

[4] W. Chen, S. Itoh and J. Shiki, On sampling in shift invariant spaces, *IEEE Trans. Information Theory,* **48** (2002), no. 10, 2802-2810.

[5] J. H. Curtiss, Introduction to functions of a complex variable, Marcel Dekker, New York, 1978.

[6] Z. Cvetkovic, and M. Vetterli, On simple oversampled A/D conversion in $L^2(R)$, *IEEE trans. Information Theory,* **47** (2001) no. 1, 146–154.

[7] I. Daubechies, and R. A. Devore, Approximating a bandlimited function using very coarsely quantized data: A family of stable sigma-delta modulators of arbitrary order, *Annals of Mathematics (2)*, **158** (2003), 679-710.

[8] G. B. Folland, Real Analysis: Modern Techniques and Their Applications, John Wiley & Sons, USA, 1999.

[9] C. S. Güntürk, Harmonic analysis of two problems in signal quantization and compression, PhD dissertation of Program in Applied and Computational Mathematics, Princeton University, Oct., 2000. [Online:] www.math.nyu.edu/~gunturk/research.html.

[10] C. S. Güntürk, Approximating a bandlimited function using very coarsely quantized data: improved error estimate in sigma-delta modulation, *Journal of American Mathematical Societity*, **17** (2004), 229-242.

[11] R. M. Gray, Spectral analysis of quantization noise in a single loop sigma-delta modulator with dc input, *IEEE Trans. Communications*, **37** (1989), 588–599.

[12] R. M. Gray, W. Chou, and P. W. Wong, Quantization noise in single-loop sigma-delta modulation with sinusoidal inputs, *IEEE Trans. Communications,* **37** (1989), no. 9, 956–967.

[13] L. K. Hua, and Y. Wang, Applications of number theory to numerical analysis, Springer-Verlag, 1981.

[14] Y. Meyer, Wavelets and Operator. Combridge University Press, 1992.

[15] H. L. Montgomery, Ten lectures on interfaces between analytic number theory and harmonic analysis, AMS, 1994.

[16] C. E. Shannon, Communications in the presence of noise, *Proc. IRE.*, **37** (1949), 10–21.

[17] E. Stein, Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals. Princeton University Press, New Jersey, 1993.

[18] N. T. Thao, Asymptotic MSE law of nth order $\Sigma\Delta$ modulators, *IEEE Trans. Circuits and Systems, Part II*, **50** (2003), 839-860.

[19] R. M. Young, An Introduction to Non-Harmonic Fourier Series: revised first edition, Academic Press, New York, 2001.