論文 / 著書情報
Article / Book Information

Title	Stability Analysis of 1-bit Sigma-Delta Modulators by Covering State Vector Transition with Hyper Cube for Specified Input Peak Amplitudes and Auto-Correlations
Authors	Mitsuhiko Yagyu, Akinori Nishihara
Citation	IEICE Trans. Fundamentals., Vol. E87-A, No. 8, pp. 1855-1862
Pub. date	2004,
URL	http://search.ieice.org/
Copyright	(c) 2004 Institute of Electronics, Information and Communication Engineers

PAPER Special Section on Digital Signal Processing

Stability Analysis of 1-Bit ΣΔ Modulators by Covering State Vector Transition with Hyper Cube for Specified Input Peak Amplitudes and Auto-Correlations

Mitsuhiko YAGYU[†], *Member and* **Akinori NISHIHARA**[††a)], *Fellow*

SUMMARY This paper presents an algorithm to analyze the stability and detect an upper-bound of every possible overload of a ΣΔ modulator for a set of input signals that are characterized by specified peak amplitudes and auto-correlations. The approach is to introduce a hyper cube in which all possible state vectors are recursively mapped into a subset of the hyper cube itself for the specified inputs and detect such a hyper cube by iteratively solving linear programming problems. Then the proposed algorithm may not identify every stable ΣΔ modulator but cannot evaluate any unstable ΣΔ modulator as a stable one. In numerical examples, two 1-bit ΣΔ modulators are analyzed, and it is revealed that a band-limitation of inputs to OSR 256 guarantees the rigorous stability even with an extension of input range to at least 240% of conventional range.

key words: *stability, oversampling ΣΔ modulator, state vector transition, optimization*

1. Introduction

High-order 1-bit sigma-delta (ΣΔ) modulators have enabled us to implement robust high-resolution A-D and D-A converters and have been applied to image signal processing, communications and sensing systems, etc. In spite of the widespread deployment of ΣΔ modulators [1], most of high-order 1-bit ΣΔ modulators have been founded on an unproven premise, that is their stability. Although many attempts to analyze the stability of ΣΔ modulators have been made [1], [2]–[7], theoretical understandings of the stability have been lagging behind due to their complicated nonlinear behavior. Therefore any methodology to design stable high-order ΣΔ modulators had never been illuminated, either.

We have started to tackle the above design problem for rigorously stable high-order ΣΔ modulators [8]–[13]. Recently we successfully designed a high-order stable 1-bit ΣΔ modulator that achieved a signal-to-noise-and-distortion ratio (SNDR) 107.6 [dB] at an oversampling ratio (OSR) 64 [12], [13]. In our design, to ensure the stability of a ΣΔ modulator shown as Fig. 1, the impulse response of the noise transfer function (NTF) $H(z)$ is constrained so as to have its l_1-norm not exceeding a specified level, and directly optimized to minimize the peak of the amplitude response of the NTF in the frequency domain. Then the quantization error $e(n)$ has strictly bounded amplitudes so that the stability is guaranteed for input amplitudes not exceeding a certain level [1], [2], [12], [13]. This condition is sufficient for the stability and may be conservative, but, due to its rigorousness and simplicity, it has been handled in the above direct optimization.

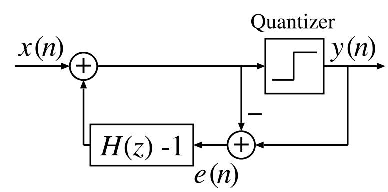


Fig. 1 1-bit ΣΔ modulator.

For the above stability, an input signal needs to be limited only on the peak amplitude and, namely, can be arbitrary within the limitation. However in some applications such as communications and audio processing etc., since input signals to a ΣΔ modulator are often band-limited, the input signals always have specific auto-correlations and thus cannot actually be arbitrary within the limitation on their amplitudes in such applications. If such specific auto-correlations of input signals can relax a tight limitation on input amplitudes and somewhat extend the permissible range of input amplitudes even with the stability, then input signals to the ΣΔ modulator can be scaled up within the permissible wider range in order to improve an SNDR of output signals and simplify an implementation of following stages. In this case, some overloads at the quantizer may occur but would be upper-bounded for a set of input signals with specified auto-correlations and peak amplitudes in the extended range. To simulate and analyze the effect of auto-correlations of narrow-band input signals, constant inputs are often used as test inputs in conventional methods, which do not rigorously guarantee the stability and hence do not upper-bound every overload, and any methods to rigorously analyze the effect have not been proposed. In this paper, we present a methodology to rigorously analyze the stability and detect an upper-bound of every possible overload with a given NTF for a set of input signals that are characterized by auto-correlations and peak amplitudes. If a time-invariant system recursively transforms a set of its all

Manuscript received November 28, 2003.
Manuscript revised March 1, 2004.
Final manuscript received May 10, 2004.

[†]The author is with the Department of Electrical and Electronic Engineering, School of Science and Technology, Meiji University, Kawasaki-shi, 214-8571 Japan.
[††]The author is with the Center for Research and Development of Educational Technology, Tokyo Institute of Technology, Tokyo, 152-8552 Japan.
a) E-mail: aki@ss.titech.ac.jp

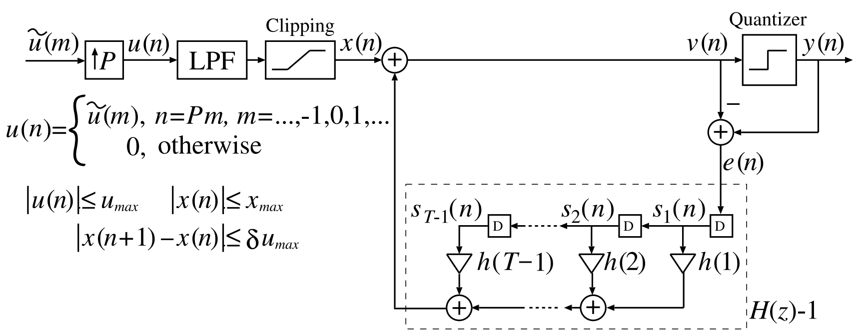


Fig. 2 A model of a ΣΔ modulator with correlated inputs for stability analysis. P stands for the OSR.

possible state vectors into a subset of the original set, the stability of the system can be proved. Our approach is to detect such a set of state vectors of a given ΣΔ modulator for a specified input peak amplitude and auto-correlation. Our method may not identify every stable ΣΔ modulator but cannot evaluate any unstable ΣΔ modulator as a stable one. Then the upper-bound of every possible overload is rigorous. In numerical examples, allowable correlated large input amplitudes of several ΣΔ modulators are analyzed and evaluated by our method.

2. Stability Analysis Method

2.1 Stability of 1-Bit ΣΔ Modulators

Let us write the Fourier Transforms of $x(n)$, $y(n)$ and $e(n)$ in Fig. 1 as $X(e^{j\omega})$, $Y(e^{j\omega})$ and $E(e^{j\omega})$, respectively. Then from Fig. 1, we obtain

$$Y(e^{j\omega}) = X(e^{j\omega}) + H(e^{j\omega})E(e^{j\omega}). \tag{1}$$

Here, we briefly define a general deterministic stability as follows; for a given set of input signals and an initial value of each internal signal, if each internal and output signal of a ΣΔ modulator can be upper-bounded by a finite number for all possible input signals of the set, then the ΣΔ modulator is defined as a stable system for the initial values and the set of input signals.

Let the impulse response of a T-tap FIR NTF $H(z)$ be $h(n)$ for $n = 0, \cdots, T-1$. From the structure shown in Fig. 1, the first sample $h(0)$ needs to be 1.0. Then a rigorous sufficient condition for the stability of a 1-bit ΣΔ modulator is known [1], [2], [12], [13] and written as

$$\text{any input peak amplitudes} \leq (3 - \|h\|_1)\Delta, \tag{2}$$

where 2Δ stands for the quantization step size. In addition, $\|h\|_1$ denotes l_1-norm of $h(n)$ and is defined as $\|h\|_1 = \sum_{n=0}^{T-1} |h(n)|$. If any input signals meet Eq. (2) and the initial amplitudes of the delay-elements of $H(z)-1$ are less than or equal to Δ, then the ΣΔ modulator is rigorously stable [1], [2], [12], [13]. In this paper, we present an algorithm to analyze the stability for given auto-correlations and input peak amplitudes that exceed Eq. (2).

2.2 Model of Correlated Inputs

Clearly, the above stability condition Eq. (2) does not limit auto-correlations of the input signals. However, if every possible input signal has specifically limited auto-correlations, then the stability may be guaranteed even for an extended range of input amplitudes. If this is true, the input signals can be properly scaled up, which may enable us to enhance an SNDR of output signals of the ΣΔ modulator and simplify implementation of following stages. First of all, in this section, we present a model of ΣΔ modulators with correlated inputs for the stability analysis.

Figure 2 illustrates a ΣΔ modulator with a given low-pass filter (LPF) for oversampling. In the figure, an input signal $\widetilde{u}(m)$ is up-sampled by a factor of P, interpolated by the LPF and clipped to a range of $\pm x_{max}$. Then every input $x(n)$ to a given ΣΔ modulator meets

$$|x(n)| \leq x_{max}. \tag{3}$$

Let the impulse response of the LPF be $h_{LPF}(n)$ for $n = 0, \cdots, T_{LPF} - 1$ and $h_{LPF}(n) = 0$ for $n < 0$ and $n \geq T_{LPF}$ for convenience. The output of the LPF can be written as

$$\sum_{k=0}^{T_{LPF}-1} h_{LPF}(k)u(n-k), \tag{4}$$

where $u(k)$ for $k = -T_{LPF} + 1, \cdots, -1$ are also defined as an initial state vector of the LPF at time 0 and meet $|u(k)| \leq u_{max}$. Then for a given $h_{LPF}(n)$, the feature of the interpolated signals can be interpreted by the convolution, i.e., certain linear equations, but the impulse response of the LPF is generally very long when the OSR is high. To simplify the analysis in this paper, we consider a simple model of the auto-correlations and then introduce

$$|x(n+1) - x(n)| \leq \delta u_{max}, \tag{5}$$

where u_{max} stands for the maximum amplitude of every possible signals $u(n)$ and $\widetilde{u}(m)$. Here, let us define a set of all possible signals of $u(n)$, which are interpolated by a factor of P and have a peak amplitude not exceeding u_{max}, as Ψ_u, and a set of all possible signals of $\widetilde{u}(m)$ as $\Psi_{\tilde{u}}$. Further, if a bounded discrete time signal $f(n)$ has a maximum amplitude, we shall write it as $\|f(n)\|_\infty = \max_n |f(n)|$. Then by using a given $h_{LPF}(n)$ and u_{max}, δu_{max} can be rewritten as

$$\delta u_{max} = \max_{u(n)\in\Psi_u} \left\| \{h_{LPF}(n-1) - h_{LPF}(n)\} * u(n) \right\|_\infty, \quad (6)$$

where $*$ denotes the convolution. Since $u(n)$ is up-sampled by a factor of P, we decompose the differential impulse response $h_{LPF}(n-1) - h_{LPF}(n)$ into P polyphase components $\tilde{h}_{LPF,\,i}(m) = h_{LPF}(Pm + i)$ for $i = 0, \cdots, P-1$ and $m = 0, \cdots, \lfloor T_{LPF}/P \rfloor$, which stands for the maximum integer not exceeding T_{LPF}/P. Then Eq. (6) can be rewritten as

$$\delta u_{max} = \max_{0\le i\le P-1} \max_{\widetilde{u}(m)\in\Psi_{\tilde{u}}} \left\| \tilde{h}_{LPF,\,i}(m) * \widetilde{u}(m) \right\|_\infty . \quad (7)$$

By introducing new independent variables $u_{i,m}$ under $|u_{i,m}| \le u_{max}$ into Eq. (7), δu_{max} can be simply obtained as

$$\delta u_{max} = \max_{0\le i\le P-1} \max_{u_{i,m}} \left| \sum_{m=0}^{\lfloor T_{LPF}/P \rfloor} \tilde{h}_{LPF,\,i}(m) u_{i,m} \right|, \quad (8)$$

$$= u_{max} \max_{0\le i\le P-1} \sum_{m=0}^{\lfloor T_{LPF}/P \rfloor} \left| \tilde{h}_{LPF,\,i}(m) \right| . \quad (9)$$

For all possible inputs $u(n)$, any $x(n)$ generated by an LPF and clipping meet Eqs. (3) and (5), but generally an arbitrary signal $x(n)$ that meets Eqs. (3) and (5) may not be actually generated with any input $\widetilde{u}(m)$. In the following, assuming that all possible $x(n)$ that meet Eqs. (3) and (5) are inputs to a given ΣΔ modulator for simplicity, we analyze the stability. The stability condition with the assumption is no longer necessary for a given LPF but still makes sense of certain sufficiency.

2.3 Maximum Amplitude of $|e(n)|$

To rigorously analyze the stability, we assume all possible $x(n)$ that meet Eqs. (3) and (5) and define $\Psi(u_{max}, x_{max}, \delta)$ as a set of those input signals $x(n)$. Then we first propose a method to obtain exact maximum of $|e(N-1)|$, which denotes the input to the loop filter $H(z) - 1$ at an arbitrary time $N-1$ in Fig. 2, for all possible signals $x(n)$ of the set $\Psi(u_{max}, x_{max}, \delta)$.

Let us write values of delay-elements of the loop-filter $H(z) - 1$ at time n as a state vector $\mathbf{s}(n)$ whose element is referred to as $s_i(n)$ for $i = 1, \cdots, T-1$. By this notation, we can write

$$s_i(n) = e(n-i). \quad (10)$$

Hereafter we express the $(T-1)$-dimensional space as $[s_1, s_2, \cdots, s_{T-1}]$. For a specified non-negative real number Δ_0, assuming all possible initial state vectors $\mathbf{s}(0)$ under

$$|s_i(0)| \le \Delta_0 \quad (11)$$

for $i = 1, \cdots, T-1$, we can write the maximization problem as

$$\begin{aligned} &\text{Maximize} && |e(N-1)| \\ &\text{Subject to} && v(n) = x(n) + \sum_{k=1}^{T-1} h(k)e(n-k), \\ &&& e(n) = \Delta\ \mathrm{sgn}[v(n)] - v(n), \\ &&& |x(n'+1) - x(n')| \le \delta u_{max}, \\ &&& |x(n)| \le x_{max} \quad \text{and} \quad |e(n'')| \le \Delta_0, \end{aligned} \quad (12)$$

where x, v, y and e are variables to be optimized for $0 \le n \le N-1$, $0 \le n' \le N-2$ and $-(T-1) \le n'' \le -1$. $h(k)$ for $k = 1, \cdots, T-1$ are the impulse response of the NTF. The above maximization problem is written with the amplitude function, which can be rearranged into two linear inequalities [14]. The *sgn* function is also nonlinear and can be rewritten as two linear functions:

$$e(n) = \begin{cases} +\Delta - v(n), & v(n) \ge 0 \\ -\Delta - v(n), & v(n) < 0 \end{cases} . \quad (13)$$

This nonlinearity prohibitively complicates the maximization problem for large N. For our approach, we first explain a way to handle the nonlinear relation between the two variables $e(0)$ and $v(0)$ for example. If we *a priori* knew whether the optimum of $v(0)$ was non-negative or not, the nonlinear relation between $e(0)$ and $v(0)$ could be rewritten as one of the two linear equations shown in Eq. (13). We propose to decompose the feasible range of $v(0)$ into two ranges $v(0) \ge 0$ and $v(0) < 0$ and obtain the maximum of $|e(N-1)|$ for each range of $v(0)$. Comparing the two maxima from those two ranges of $v(0)$, we can obtain the global maximum. By extending this, we decompose the feasible range of each variable $v(n)$ for $n = 0, \cdots, N-1$ into non-negative and negative ranges so that 2^{N-1} linear sub-problems can be derived from the nonlinear maximization problem (12). Then we can have the global optimum by comparing the maxima of all the sub-problems and, as a result, know an initial state vector and an input signal which maximize $|e(N-1)|$. Each sub-problem is formulated with fully linear equations and thus can be solved by the simplex method [14]. However the number of sub-problems to be derived exponentially increases with the growth of N, which computationally limits N.

2.4 Algorithm to Detect a Set of State Vectors Covering Every State Vector Transition

First, the initial state vector $\mathbf{s}(0)$ is assumed to be zero. We introduce a notation $\Theta_{\mathbf{s}}(n)$ as a set of all possible state vectors that can be derived at time n from the initial state vector and the set $\Psi(u_{max}, x_{max}, \delta)$. Note that l_∞-norm of every possible state vector of $\Theta_{\mathbf{s}}(n)$ is upper-bounded for a bounded n. Then if there exists a set of n_0, N' and an initial state vector such that $\Theta_{\mathbf{s}}(n_0 + N')$ is a subset of $\Theta_{\mathbf{s}}(n_0)$, the ΣΔ modulator is proved to be stable for the initial state vector and the set $\Psi(u_{max}, x_{max}, \delta)$ due to the time-invariance of the ΣΔ modulator. In other words, if arbitrary state vector of $\Theta_{\mathbf{s}}(n_0)$ changes back into any one of state vectors of $\Theta_{\mathbf{s}}(n_0)$

at time $n_0 + N'$, the $\Sigma\Delta$ modulator is stable. Then every possible state vector at time $n > n_0 + N'$ can be derived between time n_0 and time $n_0 + N'$ with a certain input signal and the initial state vector. Since there is an upper-bound of l_∞-norms of all possible state vectors derived between time n_0 and time $n_0 + N'$, every possible state vector at time $n > n_0 + N'$ is also upper-bounded by the same upper-bound, which ensures the stability. By comparing that upper-bound with the maximum l_∞-norm of all possible state vectors at time $n \le n_0$, arbitrary state vector at time $n \ge 0$ can also be upper-bounded.

However, since exact analysis and derivation of $\Theta_{\mathbf{s}}(n_0)$ at arbitrary time $n_0 > 0$ is prohibitively complicated due to the nonlinearity of the $\Sigma\Delta$ modulator, we propose to simplify $\Theta_{\mathbf{s}}(n_0)$ by a set of all possible internal and extremal points of a $(T-1)$-dimensional hyper cube which is formed by $2(T-1)$ hyper planes $s_i = \pm t$ for $i = 1, \cdots, T-1$ and then write the set as $\Theta_{cube}(t)$. In the simplification of $\Theta_{\mathbf{s}}(n_0)$, a cube $\Theta_{cube}(t_0)$ is chosen to satisfy $\Theta_{cube}(t_0) \supseteq \Theta_{\mathbf{s}}(n_0)$. Then t_0 needs to be larger than or equal to

$$l\,[\Theta_{\mathbf{s}}(n_0)] = \max_{\mathbf{s}(n_0)\in\Theta_{\mathbf{s}}(n_0)} |\mathbf{s}(n_0)|_\infty, \tag{14}$$

which stands for the maximum l_∞-norm of all possible state vectors of $\Theta_{\mathbf{s}}(n_0)$. In the simplification, first it is assumed that every state vector of $\Theta_{cube}(t_0)$ can be derived at time n_0. Then only if every possible state vector of $\Theta_{cube}(t_0)$ instead of $\Theta_{\mathbf{s}}(n_0)$ at time n_0 changes back into any one of state vectors of $\Theta_{cube}(t_0)$ at time $n_0 + N'$, the $\Sigma\Delta$ modulator is stable. In this case, that stable $\Sigma\Delta$ modulator can map all possible state vectors of $\Theta_{cube}(t_0)$ at time n_0 into a subset of $\Theta_{cube}(t_0)$ at time $n_0 + N'$. Figure 3 further explains

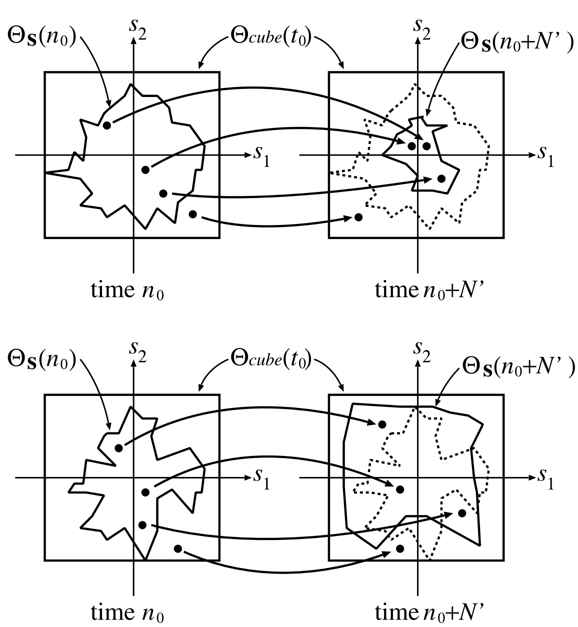


Fig. 3 An example of $\Theta_{cube}(t_0)$ that includes both $\Theta_{\mathbf{s}}(n_0)$ and $\Theta_{\mathbf{s}}(n_0+N')$, and all possible vectors of $\Theta_{cube}(t_0)$ at time n_0 are mapped into $\Theta_{cube}(t_0)$ at time $n_0 + N'$. So, if such $\Theta_{cube}(t_0)$ can be detected, though $\Theta_{\mathbf{s}}(n_0)$ may not include $\Theta_{\mathbf{s}}(n_0 + N')$, the stability can be proved.

this mapping with such $\Theta_{cube}(t_0)$ in two-dimensional case. Clearly, if $\Theta_{\mathbf{s}}(n_0 + N')$ is a subset of $\Theta_{\mathbf{s}}(n_0)$, the stability of the $\Sigma\Delta$ modulator can be proved by detecting such $\Theta_{cube}(t_0)$ including both $\Theta_{\mathbf{s}}(n_0)$ and $\Theta_{\mathbf{s}}(n_0 + N')$. However a stable $\Sigma\Delta$ modulator does not necessarily guarantee that $\Theta_{\mathbf{s}}(n_0)$ includes $\Theta_{\mathbf{s}}(n_0 + N')$. In this case, if $\Theta_{\mathbf{s}}(n_0 + N')$ is a subset of not $\Theta_{\mathbf{s}}(n_0)$ but $\Theta_{cube}(t_0)$, $\Theta_{\mathbf{s}}(n_0 + N')$ is again mapped into $\Theta_{cube}(t_0)$ at time $n > n_0 + N'$ due to the recursiveness of the mapping of $\Theta_{cube}(t_0)$ as illustrated in Fig. 3. So, if such $\Theta_{cube}(t_0)$ can be detected, that stability can also be proved. Then l_∞-norm of every state vector at time $n > n_0 + N'$ is also upper-bounded by the maximum l_∞-norm of state vectors between time n_0 and $n_0 + N'$. As $\Theta_{cube}(t_0)$ includes $\Theta_{\mathbf{s}}(n_0)$ and, namely, may have state vectors that cannot be actually derived at time n_0 from the initial state with any input signals of $\Psi(u_{max}, x_{max}, \delta)$, the above condition is no longer necessary but rigorously makes sense of the sufficiency for the stability. With this simplification, we further propose an iterative algorithm to analyze the stability and an upper-bound of l_∞-norm of every state vector at time $n > n_0 + N'$. The algorithm tries to detect such an integer N' between specified n_0 and N with the finite number of iterations and is shown below:

1. Set n_0, N and C. C stands for the maximum number of iterations carried out by this program. In addition, set the iteration counter $c := 1$.
2. For a given set $\Psi(u_{max}, x_{max}, \delta)$, analyze and calculate $l\,[\Theta_{\mathbf{s}}(n_0)]$ by maximizing $|s_i(n_0)|$ for each integer i in $1 \le i \le T-1$. To obtain each maximum of $|s_i(n_0)|$, in Eq. (12), Δ_0 is set to zero from the assumption, and $n_0 - i + 1$ is substituted into N. By solving Eq. (12), we have
$$\max_{x(n)\in\Psi} |e(n_0 - i)|. \tag{15}$$
Then $l\,[\Theta_{\mathbf{s}}(n_0)]$ is obtained as
$$\max_{1\le i\le T-1} \left\{ \max_{x(n)\in\Psi} |e(n_0 - i)| \right\} \tag{16}$$
from Eq. (10).
3. Determine t_0 as the calculated $l\,[\Theta_{\mathbf{s}}(n_0)]$ in order to simplify $\Theta_{\mathbf{s}}(n_0)$ by $\Theta_{cube}(t_0)$ which satisfies $\Theta_{cube}(t_0) \supseteq \Theta_{\mathbf{s}}(n_0)$.
4. To analyze an upper-bound of $l\,[\Theta_{\mathbf{s}}(n_0 + n)]$ for each integer n, $n = 1, \cdots, N - n_0$, set $\Delta_0 := t_0$ in Eq. (12), and then maximize $|e(n_0 + n - 1)|$ for each integer n, $n = 1, \cdots, N - n_0$. Then by solving Eq. (12), we can find the maximum $|\hat{e}(n_0 + n - 1)|$ together with a state vector at time n_0 and an input signal that generate $|\hat{e}(n_0 + n - 1)|$ for each integer n, $n = 1, \cdots, N - n_0$. Note that $|\hat{e}(n_0 + n - 1)|$ needs to be referred to only as an upper-bound of the actual exact maximum of $|e(n_0 + n - 1)|$ due to the assumption on the state vectors to be derived at time n_0.
5. Among the upper-bounds $|\hat{e}(n_0 + n - i)|$ for $1 \le i \le T-1$, the maximum is further chosen for each integer n, $n = 1, \cdots, N - n_0$, so that we can obtain an upper-bound of

$l[\Theta_s(n_0+n)]$ for each integer n, $n=1,\cdots,N-n_0$ by using Eq. (16). Note that $|\hat{e}(n_0+n-i)|$ for $n-i<0$ needs to be referred to as t_0 due to the assumption on the state vectors to be derived at time n_0.

6. Let the obtained upper-bound of $l[\Theta_s(n_0+n)]$ be $\hat{l}[t_0,\Theta_s(n_0+n)]$. Among the upper-bounds $\hat{l}[t_0,\Theta_s(n_0+n)]$, $n=1,\cdots,N-n_0$, find the upper-bound $\hat{l}[t_0,\Theta_s(n_0+n)]$ with minimum n such that $\hat{l}[t_0,\Theta_s(n_0+n)]$ is less than or equal to t_0. Then $l[\Theta_s(\tilde{n})]$ for $\forall\tilde{n}>n_0+n$ is upper-bounded by the found upper-bound, and the program indicates *Stability Proved* and stops. If all the upper-bounds $\hat{l}[t_0,\Theta_s(n_0+n)]$ for $n=1,\cdots,N-n_0$ are larger than t_0, i.e., its minimum
$$\min_{1\le n\le N-n_0}\left[\hat{l}[t_0,\Theta_s(n_0+n)]\right] \tag{17}$$
is larger than t_0, which means that the algorithm fails to prove the stability in current iteration, update t_0 by $t_0:=t_0+\mathit{\Delta}t_0$.
7. Update the iteration counter c by $c:=c+1$. If $c>C$, then the program indicates *Stability Unproved* and stops. Otherwise go to Step 4.

Any unstable ΣΔ modulators cannot have upper-bounds $\hat{l}[t_0,\Theta_s(n_0+n)]$ that satisfy $\hat{l}[t_0,\Theta_s(n_0+n)]\le t_0$ and thus cannot be identified as stable one in Step 6. However, unfortunately upper-bounds of every stable ΣΔ modulator do not necessarily satisfy the condition, which is based on the assumption to guarantee only the sufficiency for the stability. First, $\Theta_{cube}(t_0)$ does not need to be uniquely determined as shown in Step 3 and may only include $\Theta_s(n_0)$ for the sufficiency. Even if the stability of a ΣΔ modulator cannot be proved with $\Theta_{cube}(t_0)$ due to the sufficiency, it can be often proved with $\Theta_{cube}(t_0+\mathit{\Delta}t_0)$ $(\mathit{\Delta}t_0>0)$ instead of $\Theta_{cube}(t_0)$. So the algorithm updates t_0 by $t_0+\mathit{\Delta}t_0$ in Step 6 and again tries to prove the stability by using $\Theta_{cube}(t_0+\mathit{\Delta}t_0)$ in next iteration. However with the increase of t_0, the extended $\Theta_{cube}(t_0)$ naturally includes more extra state vectors that cannot be actually derived at time n_0 so that, in next iteration, each upper-bound $\hat{l}[t_0,\Theta_s(n_0+n)]$ for $n=1,\cdots,N-n_0$ is loosened, i.e., increased, which means the minimum upper-bound shown in Eq. (17) is also increased in next iteration. If the stability can be proved in next iteration, the increased minimum upper-bound in next iteration should be less than or equal to the updated t_0 so that, to prove the stability in next iteration, $t_0+\Delta t_0$ needs to be larger than the minimum upper-bound Eq. (17) in current iteration. In numerical examples, we use a slightly larger value than Eq. (17) by around 0.001Δ, where 2Δ means the quantization step size of the given ΣΔ modulator.

From the above discussion, a condition $\Theta_s(1)\subseteq\Theta_s(0)$, where $\Theta_s(0)$ is defined as $\Theta_{cube}(\Delta)$, guarantees a sufficiency for the stability and is rewritten as
$$l[\Theta_s(1)]\le\Delta. \tag{18}$$
We can derive the stability condition Eq. (2) from Eq. (18) and show it in Appendix.

3. Numerical Examples

In this section, we first analyze the stability of a 1-bit 4th-order ΣΔ modulator designed for OSR 256 by using the design algorithm [11]. The amplitude response of the NTF is illustrated in Fig. 4. The NTF has been designed so as to meet $\|h\|_1\le 2.75$ so that we have known that $|e(n)|$ is upper-bounded to Δ for arbitrary input $x(n)$ whose amplitude does not exceed 0.25Δ according to Eq. (2).

To illuminate effects of input auto-correlations, three interpolation LPFs are designed for OSRs 16, 64 and 256 as examples and incorporated as shown in Fig. 2. Their amplitude responses are depicted in Fig. 4. Then we have found that those LPFs for OSRs 16, 64 and 256 have δ in Eq. (9) as 0.28725, 0.07195 and 0.018018, respectively. Next we use $n_0=1$, $N=22$ and $C=1000$ in the proposed algorithm to keep the computing cost manageable and examine a number of input peak amplitudes $x_{max}=0.0125k\Delta$, $k=0,\cdots,80$. The proposed algorithm tries to prove the stability and an upper-bound of the peak of $|e(n)|$ for each input peak amplitude x_{max} and $u_{max}=x_{max}$. We have developped a C program and carried it out on Pentium IV 2.8GHz. Figure 5(a) shows the detected theoretical upper-bounds. From the figure, we find that, only if input signals are band-limited by the LPF for OSR 256 and have amplitudes not exceeding 0.6Δ, the ΣΔ modulator may be overloaded with $|e(n)|$ of 5.6Δ but is stable. Note that the proposed algorithm guarantees only the sufficiency for the stability. So actually the bound 0.6Δ, which is 240% of the l_1-norm-based conventional upper-bound 0.25Δ, may still be conservative, but we can see the fact that allowable range of input amplitudes can be extended to at least 0.6Δ only with the band-limitation, and the l_1-norm-based conventional upper-bound is, in fact, conservative for the band-limited input signals. In addition, we see that, even if input signals are white, actually the ΣΔ modulator can have input amplitudes within 0.28Δ with the stability. The computing times have been measured for several values of x_{max}, and we have seen that they are

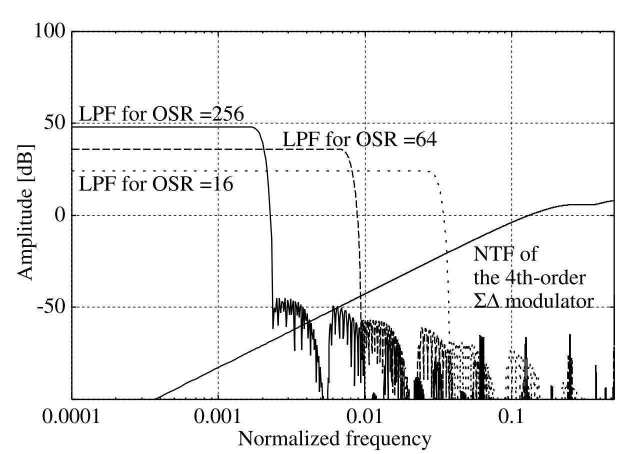
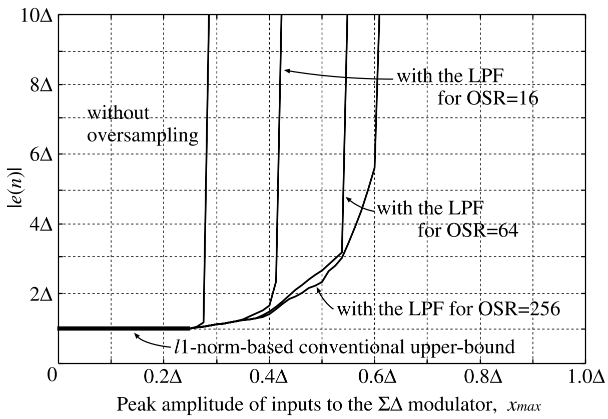
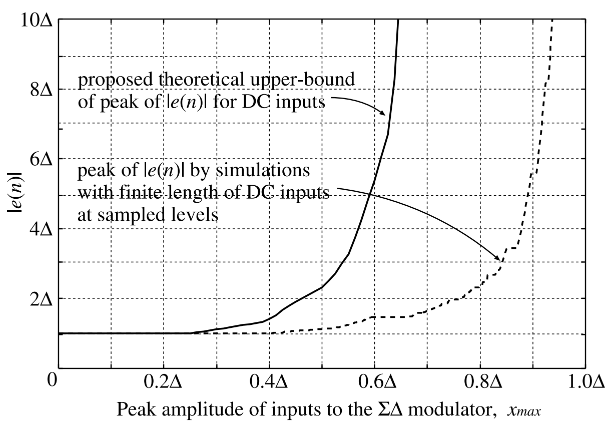


Fig. 4 Amplitude responses of the three interpolation filters and 4th-order NTF of the stability-guaranteed 1-bit ΣΔ modulator.



(a) Proposed theoretical upper-bound of the peak of $|e(n)|$ with each interpolation LPF.



(b) Proposed theoretical upper-bounds of the peak of $|e(n)|$ for DC inputs, together with the peak of $|e(n)|$ calculated by simulations with uniformly sampled input DC levels for 10^7 sampling time intervals.

Fig. 5 Proposed upper-bounds of the peak of $|e(n)|$ vs. four types of input auto-correlations and their peak amplitudes x_{max} for the 1-bit ΣΔ modulator. The design method [11] has already ensured the stability for arbitrary inputs whose amplitudes do not exceed 0.25Δ with the l_1-norm-based conventional upper-bound. The ΣΔ modulator generates two kinds of symbols; ±Δ, with which x_{max} and $|e(n)|$ can be normalized.

acceptable. For example, 44 minutes have been consumed for $x_{max} = 0.6$ and OSR 256.

The ΣΔ modulator is also examined with a special auto-correlation $\delta = 0$, i.e., DC inputs. In the same way, a number of peak DC levels $x_{max} = 0.0125k\Delta$, $k = 0, \cdots, 80$ are introduced, and then, for all possible DC inputs not exceeding each x_{max}, a theoretical upper-bound is detected and plotted in Fig. 5(b). The result guarantees that the ΣΔ modulator is stable for DC inputs not exceeding 0.64Δ. Figure 5(b) also shows the maximum of $|e(0)| \sim |e(10^7)|$ obtained by empirical simulations for each peak DC level x_{max}. The empirical peaks of $|e(n)|$ are obtained by measuring $|e(n)|$ in long time interval but may further increase with still longer time interval, and thus a rigorous upper-bound of $|e(n)|$ is undetectable only with such empirical simulations. Figures 5(a) and (b) also tell us that the stability would significantly depend on an OSR. Nevertheless, DC inputs have often been used as test inputs in conventional analysis of the stability,

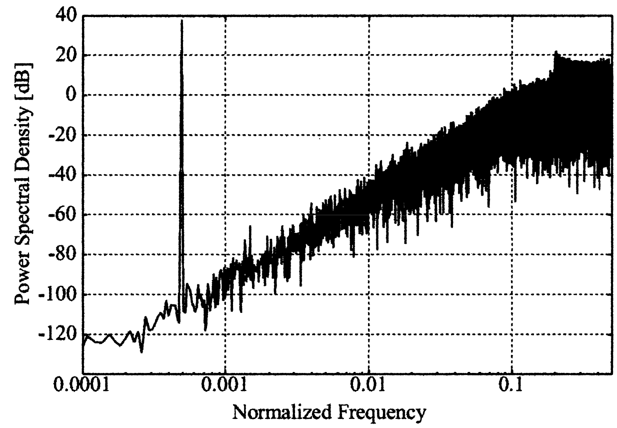
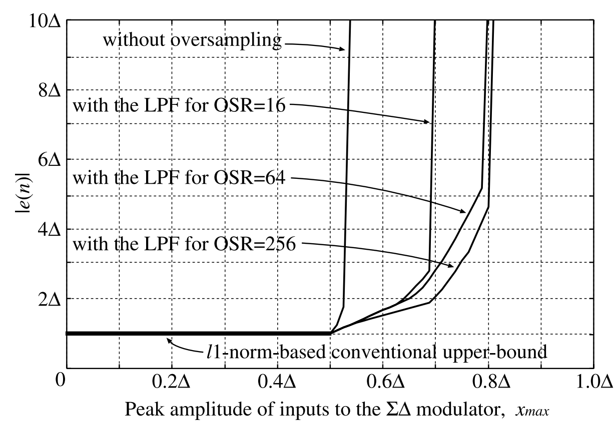
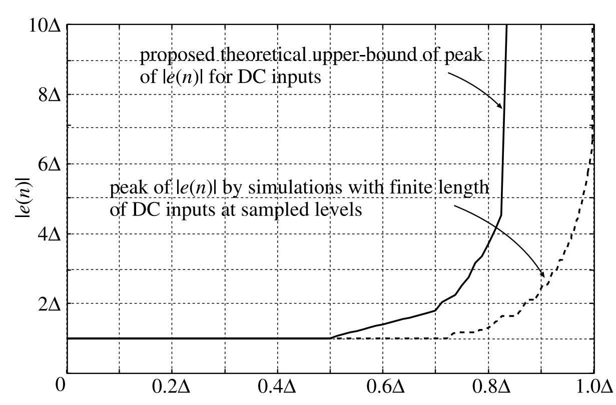


Fig. 6 Power spectral density of an output 1-bit signal.



(a) Proposed theoretical upper-bound of the peak of $|e(n)|$ with each interpolation LPF.



(b) Proposed theoretical upper-bounds of the peak of $|e(n)|$ for DC inputs, together with the peak of $|e(n)|$ calculated by simulations with uniformly sampled input DC levels for 10^7 sampling time intervals.

Fig. 7 Proposed upper-bounds of the peak of $|e(n)|$ vs. four types of input auto-correlations and their peak amplitudes x_{max} for the 1-bit ΣΔ modulator. The design method [11] has already ensured the stability for arbitrary inputs whose amplitudes do not exceed 0.5Δ with the l_1-norm-based conventional upper-bound. The ΣΔ modulator generates two kinds of symbols; ±Δ, with which x_{max} and $|e(n)|$ can be normalized.

but such analysis would not guarantee the robust design of a ΣΔ modulator for an arbitrary OSR. Of course, the proposed theoretical upper-bounds may still be conservative but rigorously ensure safe design and implementation of a ΣΔ modulator for an arbitrary OSR. A tighter theoretical upper-bound would be detected by increasing n_0 and N at the price of computing cost.

Figure 6 illustrates the power spectral density of an output 1-bit signal of this 1-bit ΣΔ modulator for an input sinusoid $0.6\Delta\sin(n\pi/1024)$, $n = 0, \cdots, 2^{16}-1$, which guarantees the stability from the experimental result shown in Fig. 5(a). Figure 6 shows that the ΣΔ modulator performs the specified noise-shaping by the NTF shown in Fig. 4. The in-band SNDR for OSR 256 is measured and obtained as 99.4 [dB].

Next another ΣΔ modulator has been designed and is analyzed for the stability. The impulse response of the NTF has been optimized as $h(0) = 1$, $h(1) = -1.25$, $h(4) = 0.25$, and $h(n) = 0$ for the other n. Thus its l_1-norm $\|h(n)\|_1 = 2.5$ ensures the stability for arbitrary inputs $x(n)$ whose amplitude does not exceed 0.5Δ. We analyze upper-bounds and the stability for correlated inputs $x(n)$ having peak amplitudes larger than 0.5Δ. Figure 7(a) shows the detected theoretical upper-bounds with the four types LPFs, and upper-bounds of $|e(n)|$ for DC inputs are plotted in Fig. 7(b). From these figures, we see that, by allowing some overloads with $|e(n)|$ of at most 4.5Δ, any input signals band-limited by the LPF for OSR 256 can be scaled up within $-0.8\Delta \sim +0.8\Delta$ with the stability. Besides, we find that the upper-bounds for DC inputs tightly bound the empirical peaks. From our experience through computer simulations, it has been confirmed that ΣΔ modulators having tightly limited $\|h(n)\|_1$ has upper-bounds of $|e(n)|$ close to empirical peaks.

4. Conclusion

This paper presents an algorithm to analyze the stability and detect an upper-bound of every possible overload of a ΣΔ modulator for a set of input signals that are characterized by specified auto-correlations and peak amplitudes. Exact analysis of all possible state vectors would be prohibitively complicated due to the nonlinearity of a ΣΔ modulator. Our approach is to introduce a hyper cube in which all possible state vectors are recursively mapped into a subset of the hyper cube itself for specified inputs and detect such a hyper cube by iteratively solving linear programming problem. Then every stable ΣΔ modulator may not be identified, but any unstable ΣΔ modulator cannot be evaluated as a stable one. In numerical examples, two 1-bit ΣΔ modulators have been analyzed, and we have found that, for the band-limited inputs with OSR 256, allowable range of the input amplitudes can be extended to at least 240% of the conventional l_1-norm-based range, and overloads may occur but are strictly upper-bounded to certain level detected by the proposed algorithm. Namely in the numerical example, it has been revealed that, with the band-limitation of inputs, the input signal power level can be significantly enhanced, which can improve an SNDR of output signals and simplify an implementation of following stages. The proposed algorithm would detect a tighter upper-bound at the price of computing cost.

References

[1] S.R. Norsworthy, R. Schreier, and G.C. Temes, Delta-Sigma Data Converters, IEEE Press, NJ, 1997.
[2] R. Schreier and Y. Yang, "Stability tests for single-bit sigma-delta modulators with second-order FIR noise transfer functions," Proc. IEEE Int. Symp. Circuits Syst., vol.3, pp.1316–1319, May 1992.
[3] S. Hein and A. Zakhor, "On the stability of sigma delta modulation," IEEE Trans. Signal Process., vol.41, no.7, pp.2322–2348, July 1993.
[4] O. Feely, "Nonlinear dynamics of sigma-delta modulation," Proc. IEEE Int. Symp. Circuits Syst., vol.6, pp.101–104, May 1994.
[5] R. Farrell and O. Feely, "Bounding the integrator outputs of second-order sigma-delta modulators," IEEE Trans. Circuits Syst. II, vol.45, no.6, pp.691–702, June 1998.
[6] R. Schreier, M.V. Goodson, and B. Zhang, "An algorithm for computing convex positively invariant sets for delta-sigma modulators," IEEE Trans. Circuits Syst. I, vol.44, no.1, pp.38–44, Jan. 1997.
[7] P. Steiner and W. Yang, "Stability of high order sigma-delta modulators," Proc. IEEE Int. Symp. Circuits Syst., vol.3, pp.52–55, May 1996.
[8] M. Yagyu and A. Nishihara, "Optimal design of oversampling quantizer by generating output signals with look-up tables," Proc. IEICE 16-th DSP Symposium, pp.463–468, Okinawa, Nov. 2001.
[9] M. Yagyu and A. Nishihara, "Stable single-bit noise-shaping quantizer based on data coding into optimized binary vectors," Proc. IEEE Int. Symp. Circuits Syst., vol.2, pp.II-384–II-387, May 2002.
[10] M. Yagyu and A. Nishihara, "Stable single-bit noise-shaping quantizer based on ΣΔ modulation and successive data coding into pre-optimized binary vectors," IEICE Trans. Fundamentals, vol.E85-A, pp.1781–1788, Aug. 2002.
[11] M. Yagyu, "Design of noise shaping FIR filters by minimizing in-band peak amplitude for stable single- and multi-bit data converters," Proc. IEEE Int. Symp. Circuits Syst., vol.1, pp.I-945–I-948, May 2003.
[12] M. Yagyu and A. Nishihara, "Design of high-order noise-shaping FIR filters for overload-free stable single- and multi-bit data converters," IEICE Trans. Fundamentals to appear.
[13] M. Yagyu and A. Nishihara, "Fast and efficient algorithm to design noise-shaping FIR filters for high-order overload-free stable sigma-delta modulators," IEEE Int. Symp. Circuits Syst., pp.I-469–472, Vancouver, May 2004.
[14] R. Fletcher, Practical methods of optimization, second ed., John Wiley & Sons, New York, 1987.

Appendix: Derivation of the Stability Condition Eq. (2) from Eq. (18)

In this section, we derive the stability condition Eq. (2) from Eq. (18).

Let us assume an arbitrary input signal $x(n)$ which satisfies $|x(n)| \leq x_{max}$. We show that the input peak amplitude x_{max} is given by Eq. (2). Since $\Theta_{\mathbf{s}}(0)$ is defined as $\Theta_{cube}(\Delta)$, we obtain $|s_i(0)| \leq \Delta$ for $i = 1, \cdots, T-1$ from Fig. 2. Note that $v(0)$ is an odd function of $x(0)$ and $s_i(0)$ for $i = 1, \cdots, T-1$ and that $h(0) = 1$ holds. Then maximum amplitude of $e(0)$ is obtained as

$$|e(0)| = |\pm\Delta - v(0)| \quad \text{(A·1)}$$

$$\leq \max_{x(0), s_i(0)} \{v(0)\} - \Delta \quad \text{(A·2)}$$

$$= \max_{x(0),s_i(0)} \left\{ x(0) + \sum_{i=1}^{T-1} h(i)s_i(0) \right\} - \Delta \quad (A\cdot 3)$$

$$= \max_{x(0)} \{x(0)\} + \sum_{i=1}^{T-1} |h(i)|\Delta - \Delta \quad (A\cdot 4)$$

$$= x_{max} + \{\|h\|_1 - 1\}\,\Delta - \Delta \quad (A\cdot 5)$$

$$= x_{max} + \{\|h\|_1 - 2\}\,\Delta. \quad (A\cdot 6)$$

From Fig. 2, $l\,[\Theta_{\mathbf{s}}(1)]$ is written as $\max_{x(0),s_i(0)} \{\Delta,\ |e(0)|\}$. By using Eq. (A· 6), Eq. (18) can be rewritten as

$$\max_{x(0),s_i(0)} |e(0)| = x_{max} + \{\|h\|_1 - 2\}\,\Delta \le \Delta. \quad (A\cdot 7)$$

Thus we obtain

$$x_{max} \le \{3 - \|h\|_1\}\,\Delta, \quad (A\cdot 8)$$

which is identical to Eq. (2).

Mitsuhiko Yagyu received the B.E., M.E. and Dr. Eng. degrees in electronics from Tokyo Institute of Technology, Tokyo, Japan, in 1993, 1995 and 1998, respectively. From 1998 to 2001, he worked for Texas Instruments Japan. He was Research Associate of the Department of Physical Electronics, Tokyo Institute of Technology in 2001 and is now Lecturer of Department of Electrical and Electronic Engineering, School of Science and Technology, Meiji University. His main research interests are in digital signal processing. He received IEICE Best Paper Award in 1999.

Akinori Nishihara received the B.E., M.E. and Dr. Eng. degrees in electronics from Tokyo Institute of Technology in 1973, 1975 and 1978, respectively. Since 1978 he has been with Tokyo Institute of Technology, where he is now Professor of the Center for Research and Development of Educational Technology. His main research interests are in one- and multi-dimensional signal processing, and its application to educational technology. He served in several editorial committees, which include an associate editor of IEEE Transactions on Circuits and Systems II from 1995 to 1997 and Editor-in-Chief of Transactions of IEICE Part A (in Japanese) from 1998 to 2000. He has been serving in IEEE Region 10 Executive Committee, as Student Activities Committee Chair (1995–1996), Treasurer (1999–2000), Educational Activities Committee Chair (2001–2004), and REP Coordinator (2003–2004). He is also an Executive Committee Member of IEEE Tokyo Section and IEEE Japan Council. He received Best Paper Awards of the IEEE Asia Pacific Conference on Circuits and Systems in 1994 and 2000, a Best Paper Award of the IEICE in 1999, and IEEE Third Millennium Medal in 2000. He also received a 4th LSI IP Design Award in 2002. Prof. Nishihara is a Fellow of IEEE and a member of EURASIP, ECS and JSET.