Frequency Domain Min-Max Optimization of Noise-Shaping Delta-Sigma Modulators

AUTHOR(S):

Nagahara, Masaaki; Yamamoto, Yutaka

CITATION:

Nagahara, Masaaki ...[et al]. Frequency Domain Min-Max Optimization of Noise-Shaping Delta-Sigma Modulators. IEEE Transactions on Signal Processing 2012, 60(6): 2828-2839

ISSUE DATE:

2012-06

URL:

http://hdl.handle.net/2433/171272

Frequency Domain Min-Max Optimization of Noise-Shaping Delta-Sigma Modulators

Masaaki Nagahara, *Member, IEEE,* Yutaka Yamamoto, *Fellow, IEEE*

***Abstract*—This paper proposes a min-max design of noise-shaping delta-sigma ($\Delta\Sigma$) modulators. We first characterize the all stabilizing loop-filters for a linearized modulator model. By this characterization, we formulate the design problem of lowpass, bandpass, and multi-band modulators as minimization of the maximum magnitude of the noise transfer function (NTF) in fixed frequency band(s). We show that this optimization minimizes the worst-case reconstruction error, and hence improves the SNR (signal-to-noise ratio) of the modulator. The optimization is reduced to an optimization with a linear matrix inequality (LMI) via the generalized KYP (Kalman-Yakubovich-Popov) lemma. The obtained NTF is an FIR (finite-impulse-response) filter, which is favorable in view of implementation. We also derive a stability condition for the nonlinear model of $\Delta\Sigma$ modulators with general quantizers including uniform ones. This condition is described as an H^∞ norm condition, which is reduced to an LMI via the KYP lemma. Design examples show advantages of our design.**

***Index Terms*—Delta-sigma modulators, min-max optimization, noise-shaping, quantization.**

I. Introduction

DELTA SIGMA ($\Delta\Sigma$, see Table I on the next page for the list of acronyms) modulators are widely used in oversampling AD (Analog-to-Digital) and DA (Digital-to-Analog) converters, by which we can achieve high performance with coarse quantizers [1], [2]. They have applications in digital signal processing systems, such as digital audio [3], [4] and digital communications [5], [6], [7]. More recently, the notion of $\Delta\Sigma$ modulators is extended to several research areas related to signal processing. In [8], [9], [10], the $\Delta\Sigma$ scheme is introduced for quantizing coefficients in finite but redundant frame expansion of signals, and is proved to outperform the standard PCM (pulse code modulation) scheme. Based on this study, $\Delta\Sigma$ scheme is also applied to compressed sensing [11], [12]. In [13], [14], dynamic quantizers as $\Delta\Sigma$ modulators are proposed for controlling linear time-invariant systems with discrete-valued control inputs. The $\Delta\Sigma$ scheme is also applied to obtain an approximate solution of large discrete quadratic programming problems [15]. For independent source separation [16] and manifold learning [17], machine learning is combined with $\Sigma\Delta$ modulation, called the $\Sigma\Delta$ *learner*.

In designing $\Delta\Sigma$ modulators, noise shaping is a fundamental issue [2]. To describe the issue of noise shaping, let us consider a general $\Delta\Sigma$ modulator shown in Fig. 1. In this figure, Q is a quantizer and $H = [H_1, H_2]$ is a linear filter with 2 inputs and 1 output. The filter H_1 shapes the signal transfer function (STF) from the input u to the output y to have a unity magnitude in the frequency band of interest. On the other hand, the filter H_2 eliminates the in-band quantization noise by shaping the noise transfer function (NTF). Then, if the input signal u is sufficiently oversampled, the frequency components in u are concentrated in the band of interest, and hence one can effectively extract the original signal u from the quantized signal y by applying a lowpass filter to y with a suitable cutoff frequency. In fact, it is theoretically shown that the reconstruction error decreases rapidly as the oversampling ratio increases [8], [18].

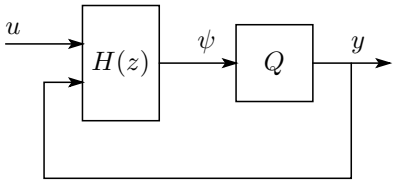

Fig. 1. $\Delta\Sigma$ modulator with loop-filter $H = [H_1, H_2]$ and quantizer Q.

A usual solution to noise shaping is to insert accumulators (or integrators) in the feedback loop to attenuate the magnitude of the NTF in low frequency. To improve upon the performance, accumulators are cascaded in various ways such as the MASH (multi-stage noise-shaping) modulators [19], [20]. This methodology is analogous to a PID (Proportional-Integral-Derivative) control [21], in which the performance of the designed system depends on the experience of the designer. That is, the conventional design is of ad hoc nature.

To obtain a systematic design method, one can adopt a more general type of transfer functions than accumulators for $H(z)$ in Fig. 1. From this point of view, the NTF zero optimization [22], [2] was proposed to shape the NTF *optimally* in the frequency band of interest, say $[0, \Omega]$. This optimization is done by changing the zeros of the NTF so as to minimize the normalized noise power given by the integral of the squared magnitude of the NTF over $[0, \Omega]$. While this method gives a systematic way to design $\Delta\Sigma$ modulators, it can yield a peak in the magnitude of the NTF at a certain frequency, since such a peak cannot be captured by an integrated or averaged objective function. A recent paper [23] has investigated this problem and proposed to use semi-infinite programming for constraining the maximum value of a function over the frequency band. This method, however, does not necessarily optimize the overall performance but only minimizes the denominator of a loop-filter. That is, the method [23] does not necessarily reduce

M. Nagahara and Y. Yamamoto are with Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan.
e-mail: nagahara@ieee.org (Nagahara), yy@i.kyoto-u.ac.jp (Yamamoto)

peaks in the NTF magnitude. Also, the computational cost for the optimization is very high due to its infinite dimensionality. Alternatively, the present authors has proposed to adopt H^∞ optimization for attenuating the NTF magnitude itself with a frequency-domain weighting function [24]. This method gives a good performance *if* a suitable weighting function was chosen. For general notion of H^∞ optimization in signal processing, see [25], [26]. The well-known Remez exchange method (aka Parks-McClellan method) [27] is related to the H^∞ optimization. The method gives a near-optimal filter that minimizes the maximum error between a given desired filter and the filter to be designed. Strictly speaking, this is not H^∞-optimal since the response is ignored on the transition frequency band.

In contrast to these methods, we propose[1] a novel design based on *min-max* optimization, which can be reduced to finite dimensional convex optimization. That is, we directly *minimize the maximum magnitude* of the NTF over the frequency band of interest. In other words, we design $\Delta\Sigma$ modulators in order to uniformly attenuate the magnitude over the prespecified band. This uniform minimization improves the *worst-case* SNR (signal-to-noise ratio) to be defined in Section III-A, of the modulator in the band of interest. Conversely, a peak of the NTF magnitude as above can deteriorate the worst-case SNR and also the dynamic range of the modulator. We propose in this paper a more effective method that does not require a selection of a weighting function.

To this end, we first characterize all stabilizing loop-filters for a linearized modulator. Then, by using this parametrization, we formulate the design problem as an optimization via a linear matrix inequality (LMI) for lowpass and bandpass modulators using the generalized Kalman-Yakubovich-Popov (KYP) lemma [30], [25]. Furthermore, we can assign arbitrarily zeros of the NTF on the unit circle in the complex plane by adding a linear matrix equality (LME) constraint to the LMI. These techniques are mostly adopted from *control theory*. Recently, control theory is effectively applied to $\Delta\Sigma$ modulator design with finite horizon predictive control [31], [32], sliding mode control [33], and robust control [34], [35], to name a few. In particular, the idea of applying the generalized KYP lemma to $\Delta\Sigma$ modulator design is proposed in [36], in which they assume a one-bit quantizer for Q and optimize the average power of the reconstruction error in low frequency for lowpass modulators. In contrast, our approach minimizes the *worst* reconstruction error, which can improve the overall SNR as mentioned above.

Stability analysis of $\Delta\Sigma$ modulators is another fundamental issue. For first-order [37] and second-order [38], [39] modulators, stability is well-studied in terms of invariant set. On the other hand, we derive a stability condition taking account of nonlinearity in $\Delta\Sigma$ modulators of arbitrary order with general quantizers including uniform ones. This condition is derived in terms of a state-space representation, and is described by the ℓ^1 norm of a linear system. This can be transformed into an H^∞-norm condition of the NTF as a sufficient condition. This H^∞-norm constraint can be equivalently expressed as an LMI via the KYP lemma [40], [41], [25]. In summary, the proposed method can be described by LMI's and LME's, which can be solved effectively by numerical optimization softwares such as YALMIP [42] and SeDuMi [43] with MATLAB.

[1]This method was first proposed in our conference articles [28], [29]. The present paper organizes these works with new results on SNR performance (Section III-A), bandpass modulator design (Section III-C), and stability theorems (Section IV). Simulation results in Section VI are also new.

TABLE I
ABBREVIATIONS

abbrev.	full name
$\Delta\Sigma$	Delta Sigma
NTF	Noise Transfer Function
STF	Signal Transfer Function
OSR	Over-Sampling Ratio
SNR	Signal-to-Noise Ratio
KYP	Kalman Yakubovich Popov
LMI	Linear Matrix Inequality
LME	Linear Matrix Equality

The organization of this paper is as follows: Section II gives characterization of all loop-filters that stabilize a linearized feedback modulator. Section III is the main section of this paper, in which we motivate the min-max design in view of SNR improvement, and then we formulate the design as a min-max optimization, which is reduced to LMI's and LME's. Section IV discusses stability of the $\Delta\Sigma$ modulator model without linearization. Section V introduces a cascade structure for high-order modulators. Section VI gives design examples to show advantages of our method. Section VII concludes our study.

Notation and Convention

Throughout this paper, we use the following notations. Abbreviations in this paper are summed up in Table I.

- $\mathcal{S}$, $\mathcal{S}'$ $\mathcal{S}$ is the set of all stable, causal, and rational transfer functions with real coefficients, and $\mathcal{S}' := \{R \in \mathcal{S} : R \text{ is strictly causal}\}$.
- ℓ^1 the Banach space of all real-valued absolutely summable sequences. For $\{v(k)\}_{k\geq 0} \in \ell^1$, the ℓ^1 norm is defined by $\|v\|_1 := \sum_{k\geq 0} |v(k)|$.
- ℓ^∞ the Banach space of all real-valued bounded sequences. For $\{v(k)\}_{k\geq 0} \in \ell^\infty$, the ℓ^∞ norm is defined by $\|v\|_\infty := \sup_{k\geq 0} |v(k)|$.
- $v * w$ convolution of two sequences $\{v(k)\}_{k\geq 0}$ and $\{w(k)\}_{k\geq 0}$, that is,

$$(v*w)(m) := \sum_{k\geq 0} v(m-k)w(k), \quad m = 0, 1, 2, \ldots.$$

For this computation, we set $v(m-k) = 0$ if $m < k$.

II. CHARACTERIZATION OF LOOP-FILTERS

In this section, we characterize all $H(z)$'s that stabilize the linearized model shown in Fig. 2. This characterization is a basis for the proposed min-max design formulated in Section III. For a stability condition taking account of the nonlinear effect of the quantizer, see the discussion in Section IV.

We first define causality, stability, well-posedness and internal stability of linear systems.

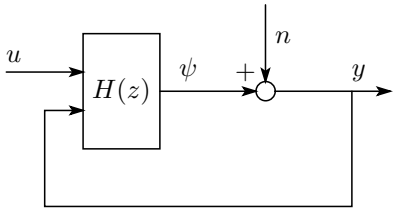

Fig. 2. Linearized model for $\Delta\Sigma$ modulator with loop-filter $H = [H_1, H_2]$.

Definition 1 (Causality and Stability): A rational transfer function $P(z)$ is said to be *(strictly) causal* if the order of the numerator of $P(z)$ is (strictly) less than that of the denominator, and said to be *stable* if the poles of $P(z)$ are all in the open unit disk $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$.

Definition 2 (Well-posedness): The feedback system in Fig. 2 is *well-posed* if there is at least one clock of delay in $H_2(z)$, that is, the transfer function $H_2(z)$ is strictly causal.

Definition 3 (Internal stability): The feedback system Fig. 2 is *internally stable* if the four transfer functions from $[u, n]^\top$ to $[\psi, y]^\top$ are all stable.

We here characterize the filter $H(z)$ that makes the linearized feedback system well-posed and internally stable. All stabilizing filters are characterized as follows:

Proposition 1: The linearized feedback system in Fig. 2 is well-posed and internally stable if and only if

$$H_1(z) = \frac{P(z)}{1+R(z)}, \quad H_2(z) = \frac{R(z)}{1+R(z)}, \\ P(z) \in \mathcal{S}, \quad R(z) \in \mathcal{S}', \tag{1}$$

where $\mathcal{S}$ denotes the set of all stable, causal, and rational transfer functions with real coefficients, and $\mathcal{S}' := \{R(z) \in \mathcal{S} : R(z) \text{ is strictly causal}\}$.

Proof: See Appendix A. ∎

By using the parameters $P(z) \in \mathcal{S}$ and $R(z) \in \mathcal{S}'$, we obtain the STF and NTF respectively as $T_{\mathrm{STF}}(z) = P(z)$ and $T_{\mathrm{NTF}}(z) = 1 + R(z)$. From this, it follows that the input/output equation of the system in Fig. 2 is given by

$$y = T_{\mathrm{STF}}\, u + T_{\mathrm{NTF}}\, n = Pu + (1+R)n. \tag{2}$$

By equation (2), the $\Delta\Sigma$ modulator can be realized by means of the design parameters $P(z) \in \mathcal{S}$ and $R(z) \in \mathcal{S}'$ as shown in Fig. 3. This structure, called *error-feedback structure* [2] or *noise-shaping coder* [1], is often applied in the digital loops required in $\Delta\Sigma$ DA converters [2]. By this block diagram, we can interpret the filter $P(z)$ as a pre-filter to shape the frequency response of the input signal, and $R(z)$ as a feedback gain for the quantization noise $Q\psi - \psi$.

III. Optimal Loop-Filter Design via Linear Matrix Inequalities and Equalities

In this section, we propose a min-max design of the loop-filter $H(z)$ by using the parametrization in Proposition 1. First, we introduce the worst-case analysis of reconstruction errors in $\Delta\Sigma$ modulators to motivate the min-max design to be proposed. We then present design procedures for lowpass and bandpass modulators.

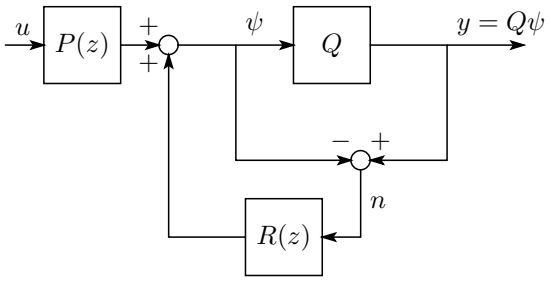

Fig. 3. Error-feedback structure of $\Delta\Sigma$ modulator with design parameters $P(z) \in \mathcal{S}$ and $R(z) \in \mathcal{S}'$.

A. Worst-case analysis of reconstruction errors

In oversampling lowpass $\Delta\Sigma$ converters with *oversampling ratio* N_{OSR} (see Table I) [2], the authors attempt to attenuate the magnitude of the NTF in the frequency band $\boldsymbol{I}_{\mathrm{B}} = [0, \Omega] \subset [0, \pi]$ where $\Omega = \pi / N_{\mathrm{OSR}}$. In a bandpass converter, the band will be $\boldsymbol{I}_{\mathrm{B}} = [\omega_0 - \Omega, \omega_0 + \Omega]$ where $\omega_0 \in (0, \pi)$ is the center frequency. We here consider a general interval $\boldsymbol{I}_{\mathrm{B}} \subset [0, \pi]$ in which the magnitude of the NTF is designed to be small. In a conventional design [22], [2], the attenuation level of the magnitude is measured by the *average* or the *root mean square*

$$\mathcal{N}_{\mathrm{average}}(T_{\mathrm{NTF}}, \boldsymbol{I}_{\mathrm{B}}) := \sqrt{\frac{1}{|\boldsymbol{I}_{\mathrm{B}}|} \int_{\boldsymbol{I}_{\mathrm{B}}} |T_{\mathrm{NTF}}(\mathrm{e}^{\mathrm{j}\omega})|^2 \mathrm{d}\omega}. \tag{3}$$

On the other hand, we consider the *worst-case* measure

$$\mathcal{N}_{\mathrm{worst}}(T_{\mathrm{NTF}}, \boldsymbol{I}_{\mathrm{B}}) := \max_{\omega \in \boldsymbol{I}_{\mathrm{B}}} \left| T_{\mathrm{NTF}}(\mathrm{e}^{\mathrm{j}\omega}) \right|. \tag{4}$$

It is easy to see that $\mathcal{N}_{\mathrm{worst}}$ gives an upper bound of $\mathcal{N}_{\mathrm{average}}$, that is,

$$\mathcal{N}_{\mathrm{average}}(T_{\mathrm{NTF}}, \boldsymbol{I}_{\mathrm{B}}) \le \mathcal{N}_{\mathrm{worst}}(T_{\mathrm{NTF}}, \boldsymbol{I}_{\mathrm{B}}).$$

Hence, minimization of $\mathcal{N}_{\mathrm{worst}}(T_{\mathrm{NTF}}, \boldsymbol{I}_{\mathrm{B}})$ leads to small $\mathcal{N}_{\mathrm{average}}(T_{\mathrm{NTF}}, \boldsymbol{I}_{\mathrm{B}})$, but not conversely. One can give an NTF with the same average $\mathcal{N}_{\mathrm{average}}$ but much larger maximum magnitude $\mathcal{N}_{\mathrm{worst}}$. That is, a small $\mathcal{N}_{\mathrm{average}}$ does not necessarily yield a small $\mathcal{N}_{\mathrm{worst}}$.

Another advantage of minimizing $\mathcal{N}_{\mathrm{worst}}$ is the worst-case optimization of the reconstruction error $y - u$ (see Fig. 3). Define the *worst-case reconstruction error* $\mathcal{E}_{\mathrm{worst}}$ by

$$\mathcal{E}_{\mathrm{worst}} := \max_{\omega \in \boldsymbol{I}_{\mathrm{B}}} \left| \hat{y}(\mathrm{e}^{\mathrm{j}\omega}) - \hat{u}(\mathrm{e}^{\mathrm{j}\omega}) \right|,$$

where $\hat{y}$ and $\hat{u}$ are, respectively, the discrete-time Fourier transforms of y and u in Fig. 3. Then this quantity can be described by the maximum magnitude $\mathcal{N}_{\mathrm{worst}}(T_{\mathrm{NTF}}, \boldsymbol{I}_{\mathrm{B}})$ of $T_{\mathrm{NTF}}(z)$ over $\boldsymbol{I}_{\mathrm{B}}$. In fact, we have the following proposition:

Proposition 2: Assume that the magnitude $|\hat{n}(\mathrm{j}\omega)|$ of the quantization noise $n = Q\psi - \psi$ is bounded on $\boldsymbol{I}_{\mathrm{B}}$, that is, there exists $C_0 > 0$ such that $\max_{\omega \in \boldsymbol{I}_{\mathrm{B}}} |\hat{n}(\mathrm{e}^{\mathrm{j}\omega})| = C_0$. Assume also that

$$|T_{\mathrm{STF}}(\mathrm{e}^{\mathrm{j}\omega})| = 1, \quad \forall \omega \in \boldsymbol{I}_{\mathrm{B}}. \tag{5}$$

Then the worst-case reconstruction error is given by

$$\mathcal{E}_{\mathrm{worst}} = C_0 \cdot \mathcal{N}_{\mathrm{worst}}(T_{\mathrm{NTF}}, \boldsymbol{I}_{\mathrm{B}}). \tag{6}$$

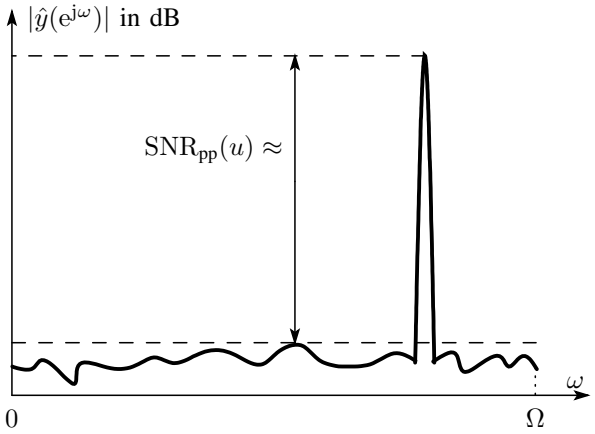

Fig. 4. Peak-to-peak SNR for a narrow-band signal.

Proof: By the relation

$$\begin{aligned}\hat{y}(e^{j\omega}) &= T_{\mathrm{STF}}(e^{j\omega})\hat{u}(e^{j\omega}) + T_{\mathrm{NTF}}(e^{j\omega})\hat{n}(e^{j\omega}) \\ &= \hat{u}(e^{j\omega}) + T_{\mathrm{NTF}}(e^{j\omega})\hat{n}(e^{j\omega}), \quad \forall \omega \in \boldsymbol{I}_{\mathrm{B}},\end{aligned}$$

we have

$$|\hat{y}(e^{j\omega}) - \hat{u}(e^{j\omega})| = |T_{\mathrm{NTF}}(e^{j\omega})\hat{n}(e^{j\omega})|, \quad \forall \omega \in \boldsymbol{I}_{\mathrm{B}}.$$

By taking the maximum over the interval $\boldsymbol{I}_{\mathrm{B}}$, we obtain (6). ■

Note that the assumption (5) holds if we choose the pre-filter $P(z)$ that has a unity magnitude response over $\boldsymbol{I}_{\mathrm{B}}$. In particular, if we take $P(z) = 1$ then we have $T_{\mathrm{STF}}(z) = 1$. By Proposition 2, optimization of $\mathcal{N}_{\mathrm{worst}}$ improves the worst-case reconstruction error $\mathcal{E}_{\mathrm{worst}}$. Minimizing $\mathcal{N}_{\mathrm{worst}}$ also improves the peak-to-peak SNR (signal-to-noise ratio) of the modulator defined by

$$\mathrm{SNR}_{\mathrm{pp}}(u) := \frac{\max_{\omega \in \boldsymbol{I}_{\mathrm{B}}} |\hat{u}(e^{j\omega})|^2}{\max_{\omega \in \boldsymbol{I}_{\mathrm{B}}} |\hat{y}(e^{j\omega}) - \hat{u}(e^{j\omega})|^2}. \tag{7}$$

Let us consider the following set of input signals:

$$\mathcal{U} := \left\{ u : \max_{\omega \in \boldsymbol{I}_{\mathrm{B}}} |\hat{u}(e^{j\omega})|^2 = 1 \right\}.$$

Suppose that the assumptions in Proposition 2 hold. Then, by Proposition 2, we have

$$\mathrm{SNR}_{\mathrm{worst}} := \min_{u \in \mathcal{U}} \mathrm{SNR}_{\mathrm{pp}}(u) = \frac{1}{C_0 \mathcal{N}_{\mathrm{worst}}(T_{\mathrm{NTF}}, \boldsymbol{I}_{\mathrm{B}})}.$$

It follows that smaller $\mathcal{N}_{\mathrm{worst}}$ leads to better *worst-case* SNR. Note that if condition (5) holds and if the input signal is sufficiently narrow-banded, $\mathrm{SNR}_{\mathrm{pp}}$ can be estimated by the difference[2] between the peak of $\hat{y}(j\omega)$ and the peak of noise, or the maximum noise level in $|\hat{y}(e^{j\omega})|$, over the frequency range $\boldsymbol{I}_{\mathrm{B}}$ (see Fig. 4).

Conversely, if $\mathcal{E}_{\mathrm{worst}}$ is as large as $\max_{\omega \in \boldsymbol{I}_{\mathrm{B}}} |\hat{u}(e^{j\omega})|$, then the $\mathrm{SNR}_{\mathrm{pp}}$ will be very poor, and the dynamic range will also be very narrow. As seen above, minimizing $\mathcal{N}_{\mathrm{average}}$ can yield a large NTF magnitude at a certain frequency, and hence the performance may be degraded. See examples in Section VI where we illustrate that minimizing $\mathcal{N}_{\mathrm{worst}}$ improves the $\mathrm{SNR}_{\mathrm{pp}}$ better than minimizing $\mathcal{N}_{\mathrm{average}}$.

[2]The difference is also known as the *spurious-free dynamic range* (SFDR).

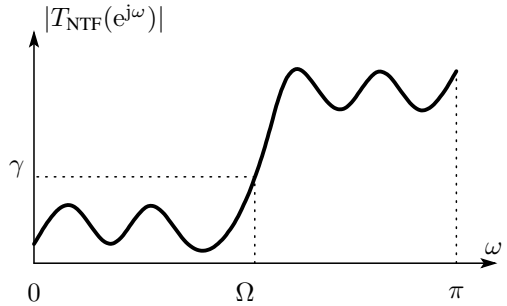

Fig. 5. Min-max optimization of the lowpass NTF in the frequency domain: minimize the maximum magnitude γ in the band $\boldsymbol{I}_{\mathrm{low}} = [0, \Omega]$.

In what follows, we set $P(z) = 1$ for simplicity, and show design methods of the loop-filter $R(z)$. Since the STF and the NTF can be designed *independently* by relation (2), one can design $P(z)$ after obtaining the loop-filter $R(z)$ such that $|P(e^{j\omega})| = 1$ over $\boldsymbol{I}_{\mathrm{B}}$ and $|P(e^{j\omega})| \ll 1$ over $[0, \pi] \setminus \boldsymbol{I}_{\mathrm{B}}$ to achieve better reconstruction performance.

B. Min-max design of lowpass modulators

We now consider the design of lowpass modulators based on the discussion given in the previous section. Our objective here is to find the loop-filter $R(z)$ that minimizes the magnitude of the frequency response of $T_{\mathrm{NTF}}(z)$ over $\boldsymbol{I}_{\mathrm{low}} := [0, \Omega]$ as shown in Fig. 5. Our problem is formulated as follows:

Problem 1 (Lowpass modulator): Given Ω $(0 < \Omega < \pi)$, find $R(z) \in \mathcal{S}'$ that solves the following min-max optimization:

$$\begin{aligned}J_{\mathrm{low}} &:= \min_{R(z) \in \mathcal{S}'} \mathcal{N}_{\mathrm{worst}}(T_{\mathrm{NTF}}, \boldsymbol{I}_{\mathrm{low}}) \\ &= \min_{R(z) \in \mathcal{S}'} \max_{\omega \in [0,\Omega]} |T_{\mathrm{NTF}}(e^{j\omega})|,\end{aligned}$$

or equivalently,

$$\begin{aligned}&\text{minimize } \gamma \text{ subject to } R(z) \in \mathcal{S}' \text{ and} \\ &\max_{\omega \in [0,\Omega]} |T_{\mathrm{NTF}}(e^{j\omega})| < \gamma.\end{aligned} \tag{8}$$

To solve this problem, we assume that $R(z)$ is a finite impulse response (FIR) filter, that is, we set

$$R(z) = \sum_{k=0}^{N} \alpha_k z^{-k}, \quad \alpha_0 = 0. \tag{9}$$

Note that the constraint $\alpha_0 = 0$ ensures $R(z) \in \mathcal{S}'$. Note also that FIR filters are often preferred to IIR filters that may cause instability attributed to quantization and recursion when they are implemented in digital devices. Therefore, the assumption to use FIR filter for $R(z)$ is not too restrictive. We then introduce a state-space realization $\{A, B, C(\alpha)\}$, such that $R(z) = C(\alpha)(zI - A)^{-1}B$, where $\alpha := [\alpha_0, \alpha_1, \ldots, \alpha_N]$,

$$A := \begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{bmatrix}, \quad B := \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \tag{10}$$

$$C(\alpha) := [\alpha_N, \alpha_{N-1}, \ldots, \alpha_1].$$

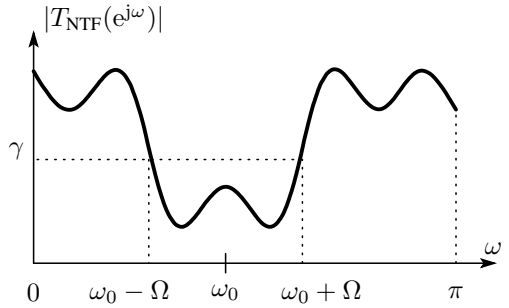

Fig. 6. Min-max optimization of the bandpass NTF in the frequency domain: minimize the maximum magnitude γ in the band $\boldsymbol{I}_{\mathrm{mid}}=[\omega_0-\Omega,\omega_0+\Omega]$.

Then inequality (8) can be described as a linear matrix inequality (LMI) by using the generalized KYP lemma [30]:

Theorem 1: Inequality (8) holds if and only if there exist symmetric matrices $Y>0$ and X such that

$$\begin{bmatrix} M_1(X,Y) & M_2(X,Y) & C(\alpha)^\top \\ M_2(X,Y)^\top & M_3(X,\gamma^2) & 1 \\ C(\alpha) & 1 & -1 \end{bmatrix} < 0, \qquad (11)$$

where

$$\begin{aligned} M_1(X,Y) &= A^\top XA + YA + A^\top Y - X - 2Y\cos\Omega, \\ M_2(X,Y) &= A^\top XB + YB, \\ M_3(X,\gamma^2) &= B^\top XB - \gamma^2. \end{aligned}$$

Proof: By the generalized KYP lemma [30, Theorem 2] for the low frequency range $\boldsymbol{I}_{\mathrm{low}}=[0,\Omega]$ in the discrete-time setting, inequality (8) is equivalent to

$$\begin{bmatrix} M_1 & M_2 \\ M_2^\top & M_3 \end{bmatrix} + \begin{bmatrix} C(\alpha) & 1 \end{bmatrix}^\top \begin{bmatrix} C(\alpha) & 1 \end{bmatrix} < 0.$$

Then applying the Schur complement [40, Sec. 2.1] to this inequality gives inequality (11). ■

By Theorem 1, the optimal coefficients $\alpha_1,\dots\alpha_N$ of the filter $R(z)$ in (9) are obtained by minimizing γ subject to (11). This LMI optimization is a convex optimization problem [40], [44], and hence can be efficiently solved by standard optimization softwares e.g., MATLAB. For optimization softwares and MATLAB codes, see Appendix C.

Remark 1: The obtained NTF $T_{\mathrm{NTF}}(z)=1+R(z)$ is an FIR filter, which is more preferred in view of implementation. On the other hand, a conventional optimal design [22], [2] yields an IIR (infinite-impulse-response) filter that has a problem of stability in digital implementation. This is an advantage of the proposed design.

C. Min-max design of bandpass modulators

Bandpass modulators are used in digital demodulation of frequency modulated analog signals, e.g., [45], [46].

We can formulate the bandpass modulator design as a min-max optimization in the same light of lowpass modulators. Fig. 6 illustrates noise shaping for bandpass modulators, where $\omega_0\in(0,\pi)$ is the center frequency and 2Ω is the bandwidth of interest. Our objective here is to minimize the magnitude of the NTF over the frequency band $\boldsymbol{I}_{\mathrm{mid}}:=[\omega_0-\Omega,\omega_0+\Omega]$. Our design process is formulated as follows:

Problem 2 (Bandpass modulator): Given $\omega_0\in(0,\pi)$ and $\Omega>0$ such that $\boldsymbol{I}_{\mathrm{mid}}=[\omega_0-\Omega,\omega_0+\Omega]\subset[0,\pi]$, find $R(z)\in\mathcal{S}'$ that solves the following min-max optimization:

$$\begin{aligned} J_{\mathrm{mid}} &:= \min_{R(z)\in\mathcal{S}'} \mathcal{N}_{\mathrm{worst}}(T_{\mathrm{NTF}},\boldsymbol{I}_{\mathrm{mid}}) \\ &= \min_{R(z)\in\mathcal{S}'} \max_{\omega\in[\omega_0-\Omega,\omega_0+\Omega]} |T_{\mathrm{NTF}}(\mathrm{e}^{\mathrm{j}\omega})|, \end{aligned}$$

or equivalently,

$$\begin{aligned} &\text{minimize } \gamma \text{ subject to } R(z)\in\mathcal{S}' \text{ and} \\ &\max_{\omega\in[\omega_0-\Omega,\omega_0+\Omega]} |T_{\mathrm{NTF}}(\mathrm{e}^{\mathrm{j}\omega})|<\gamma. \end{aligned} \qquad (12)$$

As in the lowpass modulator design, we here constrain $R(z)$ to be an FIR filter defined in (9). Let $\{A,B,C(\alpha)\}$ be state-space matrices as defined in the previous section. Then the bandpass modulator problem is also reducible to an LMI optimization via the generalized KYP lemma [30].

Theorem 2: Inequality (12) holds if and only if there exist symmetric matrices $Y>0$ and X such that

$$\begin{bmatrix} M_4(X,Y,\omega_0,\Omega) & M_5(X,Y,\omega_0) & C(\alpha)^\top \\ \overline{M}_5(X,Y,\omega_0)^\top & M_6(X,\gamma^2) & 1 \\ C(\alpha) & 1 & -1 \end{bmatrix} < 0, \qquad (13)$$

where

$$\begin{aligned} M_4(X,Y,\omega_0,\Omega) &:= A^\top XA + YA\mathrm{e}^{-\mathrm{j}\omega_0} + A^\top Y\mathrm{e}^{\mathrm{j}\omega_0} \\ &\quad - X - 2Y\cos\Omega, \\ M_5(X,Y,\omega_0) &:= A^\top XB + YB\mathrm{e}^{-\mathrm{j}\omega_0}, \\ \overline{M}_5(X,Y,\omega_0) &:= A^\top XB + YB\mathrm{e}^{\mathrm{j}\omega_0}, \\ M_6(X,\gamma^2) &:= B^\top XB - \gamma^2. \end{aligned} \qquad (14)$$

Proof: By the generalized KYP lemma [30, Theorem 2] for the mid frequency range $\boldsymbol{I}_{\mathrm{mid}}:=[\omega_0-\Omega,\omega_0+\Omega]$ in the discrete-time setting, inequality (12) is equivalent to

$$\begin{bmatrix} M_4 & M_5 \\ \overline{M}_5^\top & M_6 \end{bmatrix} + \begin{bmatrix} C(\alpha) & 1 \end{bmatrix}^\top \begin{bmatrix} C(\alpha) & 1 \end{bmatrix} < 0.$$

Then applying the Schur complement [40, Sec. 2.1] to this inequality gives inequality (13). ■

Remark 2: LMI (13) is complex-valued, however, for some LMI solvers, a real-valued LMI is required. An equivalent real-valued LMI for (13) is given in Appendix B.

Remark 3: LMI (13) with the center frequency $\omega_0=0$ is equivalent to LMI (11) for lowpass modulator. That is, Theorem 1 can be obtained as a special case of Theorem 2.

Theorem 2 can be directly extended to the following multi-band bandpass modulator design:

Problem 3 (Multi-band bandpass modulator): Given $\omega_l\in(0,\pi)$ and $\Omega_l>0$, $l=1,2,\dots,L$ such that

$$\boldsymbol{I}_l=[\omega_l-\Omega_l,\omega_l+\Omega_l]\subset[0,\pi],\ l=1,2,\dots,L,$$

find $R(z)\in\mathcal{S}'$ that solves the following min-max optimization:

$$\begin{aligned} J_{\mathrm{mb}} &:= \min_{R(z)\in\mathcal{S}'} \sum_{l=1}^{L} \mathcal{N}_{\mathrm{worst}}(T_{\mathrm{NTF}},\boldsymbol{I}_l)^2 \\ &= \min_{R(z)\in\mathcal{S}'} \sum_{l=1}^{L} \max_{\omega\in[\omega_l-\Omega_l,\omega_l+\Omega_l]} |T_{\mathrm{NTF}}(\mathrm{e}^{\mathrm{j}\omega})|^2, \end{aligned}$$

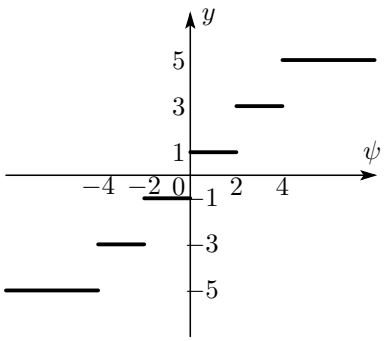

Fig. 7. Uniform quantizer Q with $M = 5$ (number of steps) and $\Delta = 2\delta = 2$ (step size).

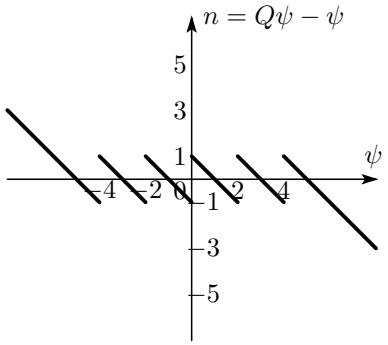

Fig. 8. Quantization error $n = Q\psi - \psi$ of the quantizer Q in Fig. 7.

or equivalently,

$$\text{minimize } \gamma_1^2 + \cdots + \gamma_L^2 \text{ subject to } R(z) \in \mathcal{S}' \text{ and } \max_{\omega \in [\omega_l - \Omega_l, \omega_l + \Omega_l]} |T_{\text{NTF}}(\mathrm{e}^{\mathrm{j}\omega})| < \gamma_l,\ l = 1, 2, \ldots, L. \quad (15)$$

Theorem 3: Inequalities (15) hold if and only if there exist symmetric matrices $Y_l > 0$ and X_l, $l = 1, 2, \ldots, L$ such that

$$\begin{bmatrix} M_4(X_l, Y_l, \omega_l, \Omega_l) & M_5(X_l, Y_l, \omega_l) & C(\alpha)^\top \\ \overline{M}_5(X_l, Y_l, \omega_l)^\top & M_6(X_l, \gamma_l^2) & 1 \\ C(\alpha) & 1 & -1 \end{bmatrix} < 0, \quad (16)$$

$$l = 1, 2, \ldots, L,$$

where M_4, M_5, $\overline{M}_5$, and M_6 are defined in (14).

Proof: A direct consequence of Theorem 2. ■

D. NTF zeros

To ensure perfect reconstruction of the DC input level, and to reduce low-frequency tones, $T_{\text{NTF}}(z)$ should have zeros at $z = 1$, or the frequency $\omega = 0$ [2]. A similar requirement is for a bandpass $\Delta\Sigma$ modulator; we set NTF zeros at a given frequency $\omega_0 \in (0, \pi)$, or $z = \mathrm{e}^{\pm \mathrm{j}\omega_0}$. The zeros of $T_{\text{NTF}}(z)$ can be assigned by linear equations (linear constraints) of $\alpha_1, \ldots, \alpha_N$. Define $\nu(z) := z^N + \sum_{k=1}^N \alpha_k z^{N-k}$. Then, $T_{\text{NTF}}(z)$ has μ zeros at $z = z_0$ if and only if

$$\left.\frac{d^k \nu(z)}{dz^k}\right|_{z=z_0} = 0, \quad k = 0, 1, \ldots, \mu - 1,$$

where $\frac{d^0 \nu(z)}{dz^0} := \nu(z)$. The LMI with these linear constraints can also be effectively solved.

IV. Stability of Nonlinear Feedback Systems

Although the linearized model in Fig. 2 is useful for analyzing and designing noise-shaping $\Delta\Sigma$ modulators as above, the stability of $\Delta\Sigma$ modulators should be analyzed with respect to their nonlinear behaviors induced by the quantizer Q. We here discuss the stability of the $\Delta\Sigma$ modulator model without linearization.

A. Stability analysis in state space

Let us first make the following assumptions:

Assumption 1: The linearized model shown in Fig. 2 is internally stable. That is, the filter $H(z) = [H_1(z), H_2(z)]$ satisfies (1).

Assumption 2: There exist real numbers $M > 0$ and $\delta > 0$ such that if $|\psi| \le M + 1$ then $|Q\psi - \psi| \le \delta$.

Note that the first assumption is necessary for the stability of the nonlinear system. The second assumption considers general quantizers including uniform ones. For example, the uniform quantizer shown in Fig. 7 has $M = 5$ and $\delta = 1$; see also Fig. 8. For uniform quantizers, the number $\Delta = 2\delta$ is called the *step size* and the interval $[-M - 1, M + 1]$ is called the *no-overload input range* [2]. Under these assumptions, we have the following lemma:

Lemma 1: Assume that Assumptions 1 and 2 hold. If $\psi(0) \le M + 1$ and if $\|p\|_1 \|u\|_\infty + \delta \|r\|_1 \le M + 1$, then we have

$$|n(k)| \le \delta, \quad |\psi(k)| \le M + 1, \quad k = 0, 1, 2, \ldots, \quad (17)$$

where p and r are respectively the impulse responses of P and R, and $\|\cdot\|_1$ and $\|\cdot\|_\infty$ denote, respectively, the ℓ^1 norm and ℓ^∞ norm of sequences.

Proof: Since the filter $H = [H_1, H_2]$ satisfies (1), we have $\psi = Pu + Rn$ where $n := Q\psi - \psi$. Then, we have $\psi(k) = (p * u)(k) + (r * n)(k)$ for $k = 0, 1, 2, \ldots$. It follows that

$$|\psi(k)| \le |(p * u)(k)| + \sum_{i=1}^k |r(i)||n(k-i)| \le \|p * u\|_\infty + \left(\max_{0 \le i \le k-1} |n(i)|\right) \sum_{i=1}^k |r(i)|.$$

If $|\psi(0)| \le M + 1$, then by Assumption 2, we have $|n(0)| = |Q\psi(0) - \psi(0)| \le \delta$, and hence

$$|\psi(1)| \le \|p * u\|_\infty + \delta \sum_{i=1}^k |r(i)| \le \|p\|_1 \|u\|_\infty + \delta \|r\|_1 \le M + 1.$$

Again by Assumption 2, we also have $|n(1)| \le \delta$. By induction on k, we deduce that $|\psi(k)| \le M + 1$ implies $|\psi(k+1)| \le M + 1$ and $|n(k+1)| \le \delta$. We thus have inequality (17). ■

This lemma gives a sufficient condition for the input ψ of the quantizer Q to be always in the no-overload range $[-M-1, M+1]$. A $\Delta\Sigma$ modulator is conventionally said to be stable if $\psi(k) \in [-M-1, M+1]$ for all $k \geq 0$ [47], [2]. However, since the modulator involves feedback, this does not necessarily guarantee boundedness of all signals in the feedback loop. To show the boundedness, we introduce a state-space model of the $\Delta\Sigma$ modulator for analyzing the stability of the feedback system.

First, invoke a minimal realization of the filter $H(z)$ be $\{A_H, [B_1, B_2], C_H, [D_H, 0]\}$, as follows:

$$H_1(z) = C_H(zI - A_H)^{-1}B_1 + D_H,$$
$$H_2(z) = C_H(zI - A_H)^{-1}B_2.$$

Then a state-space model of the closed-loop system shown in Fig. 1 is given by the following formulas:

$$\begin{aligned} x(k+1) &= A_{\rm cl}x(k) + B_u u(k) + B_n n(k), \\ n(k) &= (Q\psi - \psi)(k), \\ \psi(k) &= C_H x(k) + D_H u(k), \quad k = 0, 1, 2, \ldots, \\ A_{\rm cl} &:= A_H + B_2 C_H, \\ B_u &:= B_1 + B_2 D_H, \quad B_n := B_2. \end{aligned} \tag{18}$$

The nonlinear effect of Q is represented by the signal $n(k)$.

Consider the ideal state $x_{\rm I}(k)$, which is the state when there is no quantization, that is, when Q is identity (or $n \equiv 0$). Define the state error $e := x - x_{\rm I}$. We then have the following theorem:

Theorem 4: Suppose that the $\Delta\Sigma$ modulator shown in Fig. 1 satisfies Assumptions 1 and 2. If $\psi(0) \leq M+1$ and if

$$\|p\|_1 \|u\|_\infty + \delta \|r\|_1 \leq M + 1, \tag{19}$$

then there exists a bounded, real and monotone increasing sequence $\{\beta_k\}$ such that

$$|e(k)| \leq \beta_k, \quad k = 0, 1, 2, \ldots, \tag{20}$$

where $|e(k)|$ denotes the Euclidean norm of vector $e(k)$.

Proof: By the state-space representation (18), we have

$$\begin{aligned} x(k) &= A_{\rm cl}^k x(0) + \sum_{i=0}^{k-1} A_{\rm cl}^i B_u u(k-i) + \sum_{i=0}^{k-1} A_{\rm cl}^i B_n n(k-i) \\ &= x_{\rm I}(k) + \sum_{i=0}^{k-1} A_{\rm cl}^i B_n n(k-i). \end{aligned}$$

From this, we obtain

$$e(k) = x(k) - x_{\rm I}(k) = \sum_{i=0}^{k-1} A_{\rm cl}^i B_n n(k-i).$$

By the triangle inequality, we have

$$|e(k)| \leq \sum_{i=0}^{k-1} \|A_{\rm cl}^i B_n\| \cdot |n(k-i)|.$$

From Lemma 1, we have $|n(k)| \leq \delta$ for all $k \geq 0$. Put

$$\beta_k := \delta \sum_{i=0}^{k-1} \|A_{\rm cl}^i B_n\|.$$

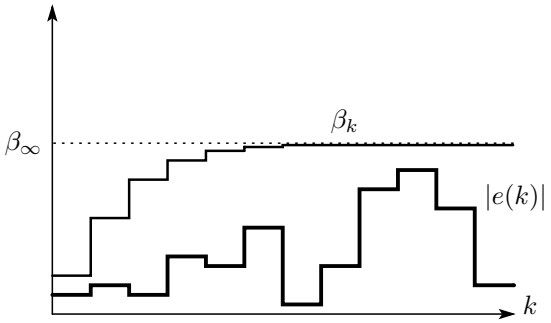

Fig. 9. Boundedness of quantization error $|e(k)|$, where β_∞ is the limiting value of $\{\beta_k\}$.

Since matrix $A_{\rm cl}$ is stable by Assumption 1, the sequence $\{\beta_k\}_{k\geq 0}$ is bounded and monotone increasing, and we have $|e(k)| \leq \beta_k$ for all $k = 0, 1, 2, \ldots$. ■

Stability condition (19) depends on the maximum amplitude of the input u. This is different from stability condition (1) for the linearized model that is independent of u. The difference is due to the nonlinearity (in particular, saturation) in the quantizer Q. Therefore, one should limit the level of inputs before it is quantized. See also the example in Section VI-A. From Theorem 4, it follows that when a $\Delta\Sigma$ modulator satisfies the condition in Theorem 4, the error $|e(k)|$ in the state space is bounded as shown in Fig. 9. As a result, the state $x(k)$ is also bounded, and we can conclude that the system is stable in a weak sense (i.e., bounded but not guaranteed to converge to zero). By Theorem 4, we derive a generalization of the stability condition given in [47] as the following corollary:

Corollary 1: Suppose that the $\Delta\Sigma$ modulator shown in Fig. 1 satisfies Assumptions 1 and 2. Define the noise-to-state transfer function $G(z)$ by $G(z) = (zI - A_{\rm cl})^{-1}B_n$, and its impulse response by g. If $\psi(0) \leq M+1$ and if inequality (19) holds, then we have $\|e\|_\infty \leq \delta \|g\|_1$.

Proof: By Theorem 4, we have

$$|e(k)| \leq \lim_{k\to\infty} \beta_k = \delta \sum_{i=0}^{\infty} \|A_{\rm cl}^i B_n\| = \delta \|g\|_1,$$

for all $k = 0, 1, 2, \ldots$. ■

B. Stability condition by an H^∞ norm inequality

Assume that $\|p\|_1 = 1$. Then, we can rewrite condition (19) in Theorem 4 as

$$\|r\|_1 \leq \frac{1}{\delta}(M + 1 - \|u\|_\infty). \tag{21}$$

By (9), we have $\|1+r\|_1 = 1 + \sum_{k=1}^N |\alpha_k| = 1 + \|r\|_1$, and we can show that (21) is equivalent to the condition given in [47], [2]:

$$\|1 + r\|_1 \leq \frac{1}{\delta}(M + 1 + \delta - \|u\|_\infty). \tag{22}$$

Let N be the order of $R(z)$. Then by the following inequality (see [48, Theorem 4.3.1]),

$$\|1 + r\|_1 \leq (2N+1)\|1 + R\|_\infty,$$

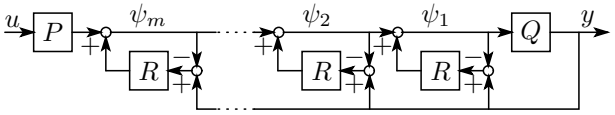

Fig. 10. Cascade of Error Feedback

we have a sufficient condition for (22):

$$\begin{aligned}\|T_{\mathrm{NTF}}\|_\infty &= \|1+R\|_\infty \\ &\le \frac{1}{(2N+1)\delta}(M+1+\delta-\|u\|_\infty).\end{aligned} \tag{23}$$

For the stability of binary $\Delta\Sigma$ modulators, the following criterion[3], called the Lee criterion, is widely used [49], [2]:

$$\|T_{\mathrm{NTF}}\|_\infty = \|1+R\|_\infty < 1.5. \tag{24}$$

From conditions (23) and (24), attenuation of the H^∞ norm of $T_{\mathrm{NTF}} = 1+R$ improves the stability. Therefore, we add the following stability constraints to the design of modulators:

$$\|T_{\mathrm{NTF}}\|_\infty = \|1+R\|_\infty < \gamma_0,$$

where $\gamma_0 > 0$ is a constant (e.g., $\gamma_0 = 1.5$ for the Lee criterion). Assuming that $R(z)$ is the FIR filter defined by (9) and also that its state-space matrices are given in (10), the above inequality is also reducible to an LMI via the KYP lemma, also known as the bounded real lemma [40], [41]:

Lemma 2: The inequality $\|T_{\mathrm{NTF}}\|_\infty < \gamma_0$ holds if and only if there exists a symmetric matrix $Z > 0$ such that

$$\begin{bmatrix} A^\top ZA - Z & A^\top ZB & C(\alpha)^\top \\ B^\top ZA & B^\top ZB - \gamma_0^2 & 1 \\ C(\alpha) & 1 & -1 \end{bmatrix} < 0.$$

Proof: The equivalence is a direct consequence of the KYP lemma (aka, bounded real lemma) [40, Sec. 2.7] and the Schur complement [40, Sec. 2.1]. ■

V. Cascade of Error Feedback for High-order Modulators

To design a high-order modulator, we can use the cascade construction of the error feedback modulators in Fig. 3. The proposed cascade structure is shown in Fig. 10. By using this structure, we have $T_{\mathrm{STF}}(z) = P(z)$ and

$$T_{\mathrm{NTF}}(z) = \big(1+R(z)\big)^m,$$

where m denotes the number of filters $R(z)$. This can be proved by the following equations:

$$\begin{aligned} \psi_m &= Pu + R(y-\psi_m), \\ y-\psi_k &= (1+R)(y-\psi_{k-1}), \quad k = m, m-1, \ldots, 2, \\ y-\psi_1 &= n. \end{aligned}$$

If $R(z) \in \mathcal{S}'$, then the linearized feedback system is stable. An advantage of this structure is that the number of taps of $R(z)$ can be reduced, and hence the implementation is much easier than a filter with a large number of taps. This structure can be applied to $\Delta\Sigma$ DA converters.

[3]Note that this is neither sufficient nor necessary for stability.

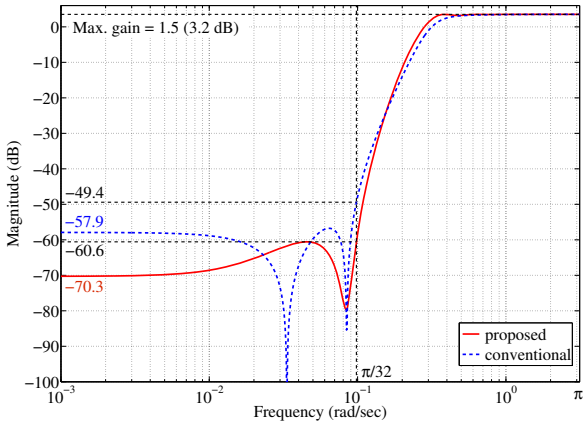

Fig. 11. NTF's: proposed (solid) and the NTF zero optimization (dash).

To satisfy the stability condition $\|T_{\mathrm{NTF}}\|_\infty < \gamma_0$, the filter $R(z)$ is designed to limit $\|1+R\|_\infty < \sqrt[m]{\gamma_0}$. If this is satisfied, we have $\|T_{\mathrm{NTF}}\|_\infty \le \|1+R\|_\infty^m < \gamma_0$, by the sub-multiplicative property of the H^∞ norm [48].

VI. Design Examples

In this section, we show two design examples of lowpass and bandpass $\Delta\Sigma$ modulators by the proposed method.

A. Lowpass modulator

We here show a design example of a high-order lowpass modulator with the cascade structure shown in Fig. 10. We set $P(z) = 1$, that is, $T_{\mathrm{STF}}(z) = 1$, and $R(z)$ be an FIR filter with 32 taps. The cutoff frequency Ω is set to be $\pi/32$. The FIR filter $R(z)$ is designed to minimize $\mathcal{N}_{\mathrm{worst}}(T_{\mathrm{NTF}}, [0,\Omega])$ defined in (4) and the coefficients are obtained by the LMI in Theorem 1, with the stability condition $\|T_{\mathrm{NTF}}\|_\infty < 1.5$, which is also described by an LMI in Lemma 2. The number m of cascades is 2, that is, the order of the modulator is $32\times 2 = 64$. We also design a modulator by the NTF zero optimization [22], [2] that minimizes the average $\mathcal{N}_{\mathrm{average}}(T_{\mathrm{NTF}}, [0,\Omega])$ defined in (3). This modulator is designed by the MATLAB function `synthesizeNTF` in the Delta-Sigma Toolbox [2], [50], where the order of T_{NTF} is 4, the over sampling ratio N_{OSR} is 32, and the stability condition $\|T_{\mathrm{NTF}}\|_\infty < 1.5$.

Fig. 11 shows the frequency responses of the proposed modulator and that by optimizing the NTF zeros. By this figure, we see that the magnitude of the proposed NTF is uniformly attenuated over $[0, \pi/32]$ while the conventional one shows peaks in this band. The difference between the two maximal magnitudes at the frequency $\omega = \pi/32$ is approximately 11.2 (dB), and the difference at low frequencies is about 12.4 (dB).

Then we run a simulation to evaluate the obtained modulators. We used MATLAB functions `simulateDSM` and `simulateSNR` in the Delta-Sigma Toolbox. Fig. 12 shows the spectrum of the output when the input is the sinusoidal wave with frequency 0.0325 (rad/sec) and amplitude 0.5. We assume a uniform quantizer with $M = 1$ and $\delta = 1/2$ (see

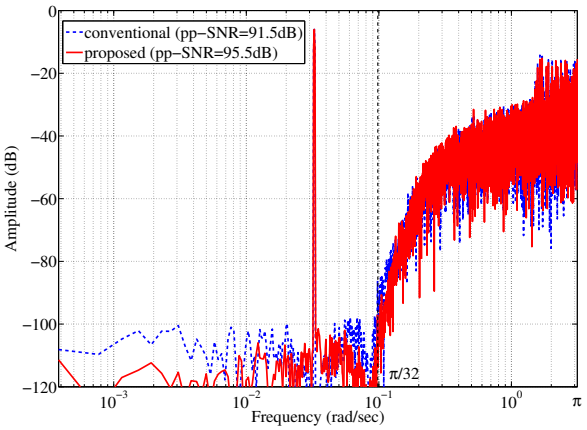

Fig. 12. The spectrum of the output of the $\Delta\Sigma$ modulators: proposed (solid) and conventional (dash), pp-SNR denotes the peak-to-peak SNR.

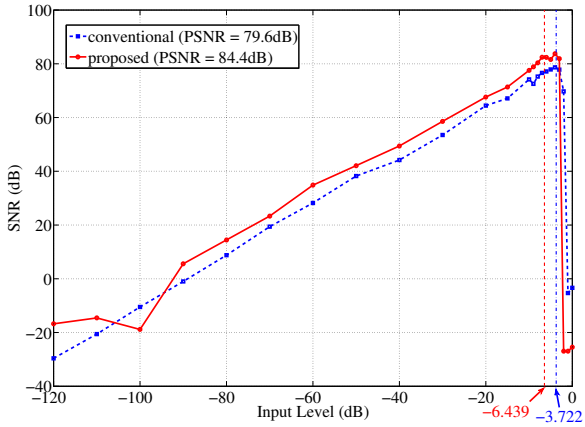

Fig. 13. The SNR versus the amplitude of the input: proposed (solid) and conventional (dash). -6.439 is the stability bound for the proposed modulator, and -3.722 is for the conventional modulator.

Assumption 2). We observe that the quantization noise is well attenuated in both cases. Note that the frequency 0.0325 (rad/sec) is taken around the first notch of the conventional NTF gain (see Fig. 11). The notch frequency is expected to give much better performance to the conventional modulator than the proposed modulator. However, the simulation shows this does not necessarily hold. In fact, the peak-to-peak SNR, $\mathrm{SNR}_{\mathrm{pp}}$ defined in (7), of our modulator is 95.5 (dB), while that of the conventional modulator is 91.5 (dB). That is, our design is superior to the conventional one in $\mathrm{SNR}_{\mathrm{pp}}$ by approximately 4.0 (dB).

Fig. 13 shows the SNR, the ratio of the signal power to the quantization noise power (SQNR), of the modulators as a function of the amplitude of the input sinusoidal wave with the frequency 0.0325 (rad/sec). For almost all amplitudes, the proposed modulator shows better performance than the conventional one, in particular, the difference of the peak SNR, or the maximum SNR is about 4.8 (dB). The figure also shows the stability bounds estimated by inequality (19) in Theorem 4. That is, the bound for the conventional modulator

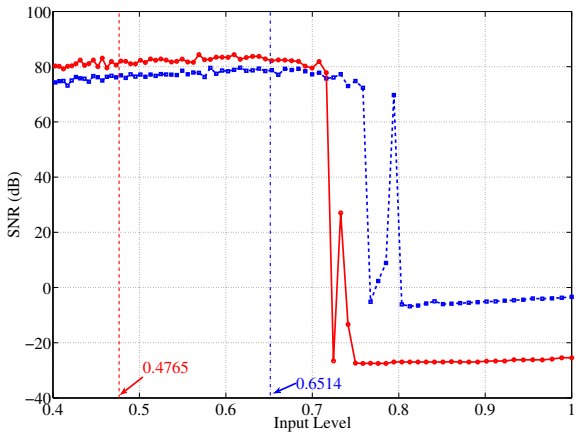

Fig. 14. Enlarged plot of Fig. 13 with linear scale for input levels.

TABLE II
COMPARISON IN FIGS. 11–13.

	max NTF (dB)	$\mathrm{SNR}_{\mathrm{pp}}$ (dB)	peak SNR (dB)
Conventional	-49.4	91.5	79.6
Proposed	-60.6	99.5	84.4
Improvement	11.2	4.0	4.8

is given by $M+1-\delta\|r\|_1 \approx 0.6514$ (-3.722 dB), and that for the proposed modulator is $M+1-\delta\|r\|_1 \approx 0.4765$ (-6.439 dB). The degradation of the SNR for high input levels is due to saturation in the quantizer that leads to instability in the modulator. We can say that if the input level is limited to the stability bound, the degradation is avoidable. We note that the conventional modulator can accept higher level of inputs. To see the difference more precisely, we show an enlarged plot in Fig. 14. The difference however does not matter if the inputs are limited to the pre-estimated bound by Theorem 4. These simulation results show that the proposed min-max (or worst-case) design gives a better SNR as mentioned in Section III-A. We summarize the results in Table II.

B. Bandpass modulator

We next show a design example of a bandpass modulator. We set $P(z) = 1$, and $R(z)$ be an FIR filter with 32 taps. The center frequency ω_0 is set to be $\pi/2$, and the bandwidth parameter Ω is $\pi/16$. The FIR filter $R(z)$ is designed by using the LMI in Theorem 2, with the stability condition $\|T_{\mathrm{NTF}}\|_\infty < 1.5$. We design two modulators, with zeros at $\omega_0 = \pm\pi/2$ and without assignment of zeros there. We also design a modulator by the NTF zero optimization [22], [2], designed by the MATLAB function `synthesizeNTF` in the Delta-Sigma Toolbox, with the order of T_{NTF} is 6, the over sampling ratio N_{OSR} is 16, the center frequency $f_0 = 1/4$, and $\|T_{\mathrm{NTF}}\|_\infty < 1.5$.

Fig. 15 shows the frequency responses of the two proposed modulators and that by optimizing the NTF zeros. We can see that the proposed modulator without assignment of zeros shows the smallest magnitude over the band $[\pi/2 - \pi/16, \pi/2 + \pi/16]$, and that of the proposed modulator with a zero at $\pi/2$ is slightly larger. To see these precisely, enlarged

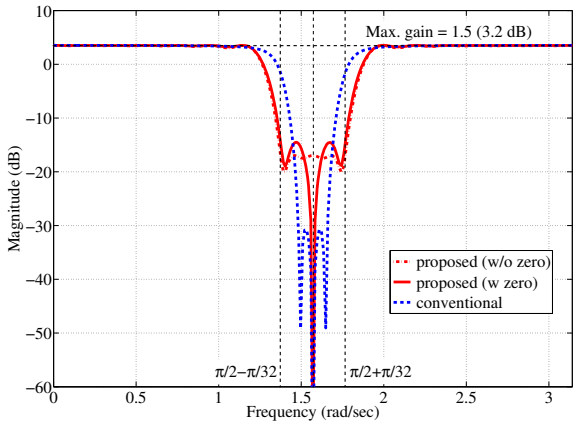

Fig. 15. Bandpass NTF's: proposed with zeros at $\omega_0 = \pm\pi/2$ (solid), proposed without assignment of zeros (dash-dots) and the NTF zero optimization (dash).

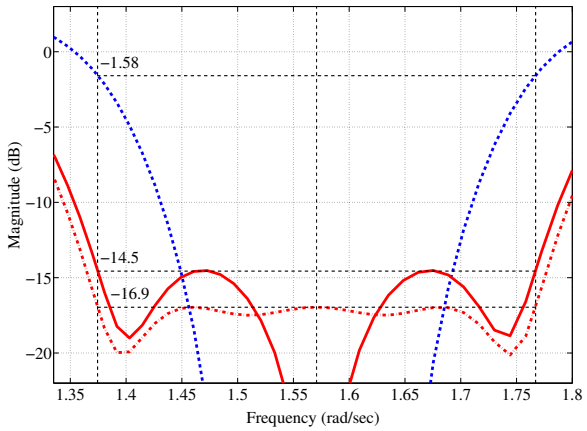

Fig. 16. Enlarged view of bandpass NTF's in Fig. 15.

figure of Fig. 15 around the center frequency is shown in Fig. 16. By this figure, the magnitudes of the proposed NTF's are uniformly attenuated over the band, while the conventional one shows a peak on the edges of the band. The differences between the magnitudes of the proposed NTF's and that of the conventional one are about 12.9 (dB) and 15.3 (dB).

Finally, we give an example of a multi-band modulator proposed in Section III-C. We set $P(z) = 1$, and $R(z)$ be an FIR filter with 32 taps. The center frequencies are set by $\omega_1 = \pi/4$, $\omega_2 = \pi/2$, and $\omega_3 = 3\pi/4$. The bandwidth parameter is $\Omega_l = \pi/16$, $l = 1, 2, 3$. We also impose the infinity norm condition $\|T_{\rm NTF}\| < 1.5$ and place zeros at ω_1, ω_2, and ω_3. Fig. 17 shows the magnitude frequency response of the NTF designed via Theorem 3. The figure shows that our design method works well.

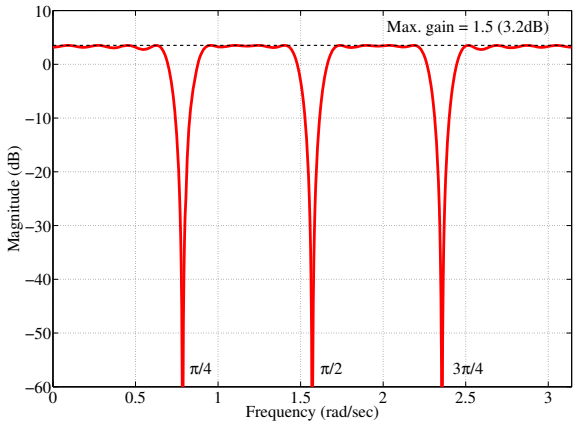

Fig. 17. Multi-band bandpass NTF designed by Theorem 3 with zeros at $\omega_1 = \pi/4$, $\omega_2 = \pi/2$, and $\omega_3 = 3\pi/4$.

VII. Conclusion

We have proposed a min-max design method of $\Delta\Sigma$ modulators. First we have characterized all stabilizing loop-filters for a linearized model. Then, based on this result, we have formulated our problem of noise shaping in the frequency domain as a min-max optimization. It is seen that the proposed min-max design has an advantage in improving SNR.

The proposed design problem is reduced to an LMI optimization, using the generalized KYP lemma, and this has a computational advantage. The assignment of NTF zeros can be taken care of by an LME. We have given a stability analysis of the $\Delta\Sigma$ modulator model without linearization and derived an H^∞-norm condition for stability, which is also described as an LMI via the KYP lemma. The obtained NTF is an FIR filter, which is favorable from the implementation viewpoint. Design examples have shown effectiveness of our method.

Future work includes STF optimization as in [23], or adaptive quantization as in [51] combined with the proposed optimal filter.

Acknowledgments

This research is supported in part by the JSPS Grant-in-Aid for Scientific Research (B) No. 2136020318360203, Grant-in-Aid for Exploratory Research No. 22656095, and the MEXT Grant-in-Aid for Young Scientists (B) No. 22760317.

Appendix

A. *Proof of Proposition 1*

In this proof, we adopt a standard technique of control theory [52].

First assume that $H_1(z)$ and $H_2(z)$ are given by (1) for some $P(z) \in \mathcal{S}$ and $R(z) \in \mathcal{S}'$. Since $R(z) \in \mathcal{S}'$, $R(z)$ is strictly causal and so is $H_2(z) = R(z)/(1 + R(z))$. This implies that the system is well-posed. For internal stability, we need to show that the four transfer functions $1/(1 - H_2(z))$, $H_1(z)/(1-H_2(z))$, and $H_2(z)/(1-H_2(z))$ are all stable (i.e., their poles are inside the unit circle in the complex plane). By the equalities in (1), we have $1/(1 - H_2(z)) = 1 + R(z) \in \mathcal{S}$, and hence $H_1(z)/(1 - H_2(z)) = P(z)$ and $H_2(z)/(1 - H_2(z)) = R(z)$ are stable.

Next assume that the feedback system is well-posed and internally stable. Define $R := H_2/(1 - H_2)$ and $P := H_1/(1 - H_2)$. Since $H_2(z)$ is strictly proper from the well-posedness, so is $R(z)$. Then by the internal stability of the

feedback system, $R = H_2/(1-H_2)$ and $P = H_1/(1-H_1)$ are stable, that is $R(z) \in \mathcal{S}'$ and $P(z) \in \mathcal{S}$. ■

B. Real-valued LMI for Theorem 2

For a Hermitian matrix $F \in \mathbb{C}^{n\times n}$ the inequality $F < 0$ is equivalent to ([44])

$$\begin{bmatrix} \mathrm{Re}\, F & -\mathrm{Im}\, F \\ \mathrm{Im}\, F & \mathrm{Re}\, F \end{bmatrix} < 0.$$

Hence we obtain the following real-valued LMI for (13):

$$\begin{bmatrix} M_r(X,Y,\alpha) & -M_i(Y) \\ M_i(Y) & M_r(X,Y,\alpha) \end{bmatrix} < 0,$$

where

$$\begin{aligned}
M_r(X,Y,\alpha) &:= \begin{bmatrix} M_{r1}(X,Y) & M_{r2}(X,Y) & C(\alpha)^\top \\ M_{r2}(X,Y)^\top & M_{r3}(X,\gamma) & 1 \\ C(\alpha) & 1 & -1 \end{bmatrix}, \\
M_{r1}(X,Y) &:= A^\top XA + (A^\top Y + YA)\cos\omega_0 \\
&\quad - X - 2Y\cos\Omega, \\
M_{r2}(X,Y) &:= A^\top XB + YB\cos\omega_0, \\
M_{r3}(X,\gamma) &:= B^\top XB - \gamma^2, \\
M_i(Y) &:= \begin{bmatrix} M_{i1}(Y) & M_{i2}(Y) & 0 \\ -M_{i2}^\top & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \\
M_{i1}(Y) &:= (A^\top Y - YA)\sin\omega_0, \\
M_{i2}(Y) &:= -YB\sin\omega_0.
\end{aligned}$$

■

C. MATLAB codes for optimal NTF

We here introduce MATLAB codes for executing numerical computation of the design proposed in this paper. The codes are downloadable from the following web site:

`http://www-ics.acs.i.kyoto-u.ac.jp/~nagahara/ds/`

This site provides a MATLAB function `NTF_MINMAX`, which is the main function to design optimal modulators. Note also that to execute the codes in this section, Control System Toolbox [53], YALMIP [42], and SeDuMi [43] are needed. We use Control System Toolbox for defining state-space representation of systems. YALMIP is a parser for LMI description and SeDuMi is a solver for convex optimization problem including LMI's with the self-dual embedding technique. This function computes the optimal NTF and $R(z)$ minimizing $\gamma > 0$ subject to LMI (11) for lowpass modulators and (13) for bandpass modulators. The H^∞-norm condition of the NTF and assignment of the NTF zeros can be also included using Lemma 2.

For example, the optimal lowpass NTF shown in Section VI-A is obtained by

```
[ntf2,R]=NTF_MINMAX(32,32,1.5^(1/2),0,0);
ntf=ntf2^2;
```

The optimal bandpass NTF with zeros at $z = \mathrm{e}^{\pm \mathrm{j}\pi/2}$ shown in Section VI-B is obtained by

```
[ntf,R]=NTF_MINMAX(32,16,1.5,1/4,1);
```

For the optimal multi-band bandpass NTF shown in Section VI-B is also obtained by using another function `NTF_MINMAX_MB` as

```
ff=[1/8,1/4,3/8];
[ntf,R]=NTF_MINMAX_MB(32,64,1.5,ff,1);
```

Remark 4: When one runs the codes, a message "`Run into numerical problems`" may appear. This means that there was some kind of a numerical problem encountered in optimization, and the usefulness of the returned solution should be judged by the designer. This may happen occasionally in numerical LMI optimization. For example, in numerical optimization with an LMI condition $M > 0$, the minimum eigenvalue of M may be slightly negative due to numerical problems. In many cases, this does not matter. To avoid this, one can adopt very small $\varepsilon > 0$ and rewrite $M > 0$ as $M > \varepsilon I$.

REFERENCES

[1] S. R. Norsworthy, R. Schreier, and G. C. Temes, *Delta-Sigma Data Converters*. IEEE Press, 1997.
[2] R. Schreier and G. C. Temes, *Understanding Delta-Sigma Data Converters*. Wiley Interscience, 2005.
[3] E. Janssen and D. Reefman, "Super-audio CD: An introduction," *IEEE Signal Process. Mag.*, vol. 20, no. 4, pp. 83–90, 2003.
[4] U. Zölzer, *Digital Audio Signal Processing*, 2nd ed. John Wiley & Sons, 2008.
[5] K. Vleugels, S. Rabii, and B. A. Wooley, "A 2.5-V sigma-delta modulator for broadband communicatinos applications," *IEEE J. Solid-State Circuits*, vol. 36, no. 12, pp. 1887–1899, 2001.
[6] A. Rusu, B. Dong, and M. Ismail, "Putting the "FLEX" in flexible mobile wireless radios," *IEEE Circuits Devices Mag.*, vol. 22, no. 6, pp. 24–30, 2006.
[7] U. Gustavsson, T. Eriksson, and C. Fager, "Quantization noise minimization in ΣΔ modulation based RF transmitter archtectures," *IEEE Trans. Circuits Syst. I*, vol. 57, no. 12, pp. 3082–3091, Dec. 2010.
[8] I. Daubechies and R. DeVore, "Approximating a bandlimited function using very coarsely quantized data: A family of stable sigma-delta modulators of arbitrary order," *Annals of Mathematics*, vol. 158, no. 2, pp. 679–710, 2003.
[9] J. Benedetto, A. Powell, and O. Yılmaz, "Sigma-delta quantization and finite frames," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1990–2005, May 2006.
[10] M. Lammers, A. M. Powell, and O. Yılmaz, "Alternative dual frames for digital-to-analog conversion in sigma–delta quantization," *Adv. Comput. Math.*, vol. 32, no. 1, pp. 73–102, Jan. 2010.
[11] P. Boufounos and R. G. Baraniuk, "Sigma delta quantization for compressive sensing," in *Proc. SPIE Waveletts XII*, vol. 6701, Aug. 2007.
[12] C. S. Güntürk, M. Lammers, A. Powell, R. Saab, and O. Yılmaz, "Sigma delta quantization for compressed sensing," in *44th Annual Conf. on Inf. Sci. and Syst. (CISS)*, 2010, pp. 1–6.
[13] S. Azuma and T. Sugie, "Optimal dynamic quantizers for discrete-valued input control," *Automatica*, vol. 44, no. 2, pp. 396–406, Feb. 2008.
[14] ——, "Synthesis of optimal dynamic quantizers for discrete-valued input control," *IEEE Trans. Autom. Control*, vol. 53, no. 9, pp. 2064–2075, Oct. 2008.
[15] S. Callegari, F. Bizzarri, R. Rovatti, and G. Setti, "On the approximate solution of a class of large discrete quadratic programming problems by ΔΣ modulation: The case of circulant quadratic forms," *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 6126–6139, 2010.
[16] A. Fazel, A. Gore, and S. Chakrabartty, "Resolution enhancement in ΣΔ learners for superresolution source separation," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1193–1204, Mar. 2010.
[17] A. Gore and S. Chakrabartty, "A min–max optimization framework for designing ΣΔ learners: Theory and hardware," *IEEE Trans. Circuits Syst. I*, vol. 57, no. 3, pp. 604–617, Mar. 2010.
[18] F. Krahmer and R. Ward, "Lower bounds for the error decay incurred by coarse quantization schemes," *Applied and Computational Harmonic Analysis*, vol. 32, no. 1, pp. 131–138, Jan. 2012.

[19] T. Hayashi, Y. Inabe, K. Uchimura, and A. Iwata, "A multistage delta-sigma modulator without double integration loop," *ISSCC Digest of Technical Papers*, pp. 182–183, 1986.
[20] J. C. Candy and A. Huynh, "Double integration for digital-to-analog conversion," *IEEE Trans. Commun.*, vol. 34, no. 12, pp. 1746–1756, 1986.
[21] A. Datta, M.-T. Ho, and S. P. Bhattacharyya, *Structure and Synthesis of PID Controllers*. Springer, 1995.
[22] R. Schreier, "An empirical study of high-order single-bit delta-sigma modulators," *IEEE Trans. Circuits Syst. II*, vol. 40, no. 8, pp. 461–466, 1993.
[23] C.-F. Ho, B. Ling, J. Reiss, Y.-Q. Liu, and K.-L. Teo, "Design of interpolative sigma delta modulators via semi-infinite programming," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 4047–4051, Oct. 2006.
[24] M. Nagahara, T. Wada, and Y. Yamamoto, "Design of oversampling delta-sigma DA converters via H^∞ optimization," in *Proc. of IEEE ICASSP*, vol. III, 2006, pp. 612–615.
[25] M. Nagahara, "Min-max design of FIR digital filters by semidefinite programming," in *Applications of Digital Signal Processing*. InTech, Nov. 2011, pp. 193–210.
[26] Y. Yamamoto, M. Nagahara, and P. P. Khargonekar, "Signal reconstruction via H^∞ sampled-data control theory—Beyond the Shannon paradigm," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 613–625, Feb. 2012.
[27] T. Parks and J. McClellan, "Chebyshev approximation for nonrecursive digital filters with linear phase," *IEEE Trans. Circuit Theory*, vol. 19, no. 2, pp. 189–194, Mar. 1972.
[28] M. Nagahara and Y. Yamamoto, "Optimal noise shaping in delta-sigma modulators via generalized KYP lemma," in *Proc. of IEEE ICASSP*, 2009, pp. 3381–3384.
[29] ——, "Optimal design of delta-sigma modulators via generalized KYP lemma," in *Proc. of ICROS-SICE International Joint Conference*, 2009, pp. 4376–4379.
[30] T. Iwasaki and S. Hara, "Generalized KYP lemma: unified frequency domain inequalities with design applications," *IEEE Trans. Autom. Control*, vol. AC-50, pp. 41–59, 2005.
[31] D. Quevedo and G. Goodwin, "Multistep optimal analog-to-digital conversion," *IEEE Trans. Circuits Syst. I*, vol. 52, no. 3, pp. 503–515, Mar. 2005.
[32] J. Østergaard, D. Quevedo, and J. Jensen, "Real-time perceptual moving-horizon multiple-description audio coding," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4286–4299, Sep. 2011.
[33] S.-H. Yu, "Analysis and design of single-bit sigma-delta modulators using the theory of sliding modes," *IEEE Trans. Control Syst. Technol.*, vol. 14, no. 2, pp. 336–345, Mar. 2006.
[34] F. Yang and M. Gani, "An H_∞ approach for robust calibration of cascaded sigma–delta modulators," *IEEE Trans. Circuits Syst. I*, vol. 55, no. 2, pp. 625–634, Mar. 2008.
[35] J. McKernan, M. Gani, F. Yang, and D. Henrion, "Optimal low-frequency filter design for uncertain 2-1 sigma-delta modulators," *IEEE Signal Process. Lett.*, vol. 16, no. 5, pp. 362–365, May 2009.
[36] M. Osqui, M. Roozbehani, and A. Megretski, "Semidefinite programming in analysis and optimization of performance of sigma-delta modulators for low frequencies," in *Proc. of the American Control Conf.*, 2007, pp. 3582–3587.
[37] A. Teplinsky, E. Condon, and O. Feely, "Driven interval shift dynamics in sigma-delta modulators and phase-locked loops," *IEEE Trans. Circuits Syst. I*, vol. 52, no. 6, pp. 1224–1235, Jun. 2005.
[38] C. S. Güntürk and N. T. Thao, "Refined error analysis in second-order $\Sigma\Delta$ modulation with constant inputs," *IEEE Trans. Inf. Theory*, vol. 50, no. 5, pp. 839–860, May 2004.
[39] C. Y.-F. Ho, B. W.-K. Ling, J. D. Reiss, and X. Yu, "Global stability, limit cycles and chaotic behaviors of second order interpolative sigma delta modulators," *International Journal of Bifurcation and Chaos*, vol. 21, no. 6, pp. 1755–1772, 2011.
[40] S. Boyd, L. E. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. SIAM, 1994.
[41] Y. Yamamoto, B. D. O. Anderson, M. Nagahara, and Y. Koyanagi, "Optimizing FIR approximation for discrete-time IIR filters," *IEEE Signal Process. Lett.*, vol. Vol. 10, No. 9, 2003.
[42] J. Löfberg, "YALMIP : A toolbox for modeling and optimization in MATLAB," in *Proc. IEEE International Symposium on Computer Aided Control Systems Design*, 2004, pp. 284–289. [Online]. Available: http://users.isy.liu.se/johanl/yalmip/
[43] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," 2001. [Online]. Available: http://sedumi.ie.lehigh.edu/
[44] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
[45] S. Jantzi, R. Schreier, and M. Snelgrove, "Bandpass sigma-delta analog-to-digital conversion," *IEEE Trans. Circuits Syst.*, vol. 38, no. 11, pp. 1406–1409, Nov. 1991.
[46] S. Jantzi, K. Martin, and A. Sedra, "Quadrature bandpass $\Delta\Sigma$ modulation for digital radio," *IEEE J. Solid-State Circuits*, vol. 32, no. 12, pp. 1935–1950, Dec. 1997.
[47] J. G. Kenney and L. R. Carley, "Design of multibit noise-shaping data converters," *Analog Int. Circuits Signal Processing Journal*, vol. Vol. 3, pp. 259–272, 1993.
[48] M. A. Dahleh and I. J. Diaz-Bobillo, *Control of Uncertain Systems*. Prentice Hall, 1995.
[49] K. C. H. Chao, S. Nadeem, W. L. Lee, and C. G. Sodini, "A higher order topology for interpolative modulators for oversampling A/D conversion," *IEEE Trans. Circuits Syst.*, vol. 37, no. 3, pp. 309–318, 1990.
[50] R. Schreier, "Delta sigma toolbox." [Online]. Available: http://www.mathworks.com/matlabcentral/fileexchange/19
[51] J. Østergaard and R. Zamir, "Multiple-description coding by dithered delta-sigma quantization," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4661–4675, Oct. 2009.
[52] J. C. Doyle, B. A. Francis, and A. R. Tannenbaum, *Feedback Control Theory*. Maxwell Macmillan, 1992.
[53] Mathworks, "Control system toolbox users guide," 2010. [Online]. Available: http://www.mathworks.com/products/control/

Masaaki Nagahara (S'00–M'03) received the Bachelor's degree in engineering from Kobe University in 1998, the Master's degree and the Doctoral degree in informatics from Kyoto University in 2000 and 2003. He is currently an Assistant Professor at the Graduate School of Informatics, Kyoto University. His research interests include digital signal processing and digital control systems. He was a recipient of Outstanding Paper Awards from the Society of Instrument and Control Engineers (SICE) in 1999. He is a member of IEEE, ISCIE, SICE, and IEICE.

Yutaka Yamamoto (M'83–SM'93–F'98) received the Ph.D. degree in mathematics from the University of Florida, Gainesville, in 1978, under the guidance of Professor Rudolf E. Kalman. He is currently a Professor in the Department of Applied Analysis and Complex Dynamical Systems, Graduate School of Informatics, Kyoto University, Kyoto, Japan.

His current research interests include the theory of sampled data control systems, its application to digital signal processing, realization and robust control of distributed parameter systems and repetitive control.

Dr. Yamamoto was the recipient of G. S. Axelby Outstanding Paper Award of the IEEE Control Systems Society in 1996, the Commendation for Science and Technology by the Minister of Education, Culture, Sports, Science and Technology, Prize for Science and Technology, in Research Category in 2007, Distinguished Member Award of the Control Systems Society of the IEEE in 2009, and several other awards from the Society of Instrument and Control Engineers (SICE) and the Institute of Systems, Control and Information Engineers (ISCIE). He has served as Senior Editor of the Trans. Automatic Control (TAC) for 2010–2011, and as an AE for several journals including TAC, Automatica, MCSS, SCL. He was a chair of the Steering Committee of MTNS for 2006–2008, and the General Chair of the MTNS 2006 in Kyoto.

He is currently President-Elect of the Control Systems Society (CSS) of the IEEE, and was a vice president for 2005–2008 of CSS, and President of ISCIE of Japan for 2008–2009. He is a fellow of the IEEE and SICE, Japan.