DELTA-SIGMA MODULATORS FOR CONSTANT ENVELOP TRANSMISSIONS WITH GUARANTEED STABILITY

Shuichi Ohno

Hiroshima University, 1-4-1 Kagamiyama, Higashi-Hiroshima, 739-8527, JAPAN

ABSTRACT

1-bit Delta-Sigma-based transmitters enable constant envelop transmissions. This paper synthesizes a stable 1-bit Delta-Sigma modulator that minimizes the mean squared quantization error at the output of a Delta-Sigma-based transmitter. The 1-bit Delta-Sigma modulator is designed by solving convex optimization problems, which can be solved numerically. A numerical example is provided to demonstrate our synthesis.

Index Terms— Delta-Sigma modulator, quantizer, constant envelop

1. INTRODUCTION

Efficient power amplifiers (PAs) are essential for modern wireless communications. For example, the massive multiple-input multiple-output (MIMO) system, which has been attracting much attention as one of the technologies for 5G wireless communication systems [1], utilizes a large number of antennas and then requires a large number of radio-frequency (RF) chains including power amplifiers. Reasonable amplifiers are necessary for the deployment of massive MIMO systems [2].

To capture the variation of a signal, the peak to average power ratio (PAPR) of a signal is defined as the peak power divided by its average power. For signals with large PAPR, the efficiency of the power amplifier is low, since the power amplifier has to be operated in a backed off. Moreover, the power amplifier should be linear up to the maximum input power to avoid distortion. The large backoff and the linearity increase the complexity of the power amplifier, which also increases its implementation cost.

After encoding an analog signal to a digital signal, one may utilize phase shift-keying (PSK), which is well-known as a constant envelop transmission. However, the high-order PSK modulation exhibits a poor bit error rate (BER) performance. Another possible remedy is the transmission with Delta-Sigma ($\Delta\Sigma$) modulators [3], where the information signal is encoded to a signal which takes only a few discrete values. If the signal is encoded to a bi-level signal, the power amplifier works at its maximum efficiency without any power backoff.

In a 1-bit $\Delta\Sigma$ modulator, which consists of a 1-bit uniform quantizer and a feedback filter, the information signal is oversampled and is fed back. The quantization error of the 1-bit quantizer is processed with an error feedback filter and then the filtered error is added to the input signal, which is fed into the 1-bit quantizer. (see Fig. 2 in Section II). Then, the output of the $\Delta\Sigma$ modulator is carrier-modulated and is amplified. Finally, the signal is filtered by a bandpass filter to reconstruct the information signal and to reduce the effect of the quantization error. Since $\Delta\Sigma$ modulators can be implemented digitally, all-digital transmitter is possible with $\Delta\Sigma$-based transmitters [4, 5], which is also suitable to realize software defined radios.

One of the main disadvantages of $\Delta\Sigma$-based transmitters is the quantization noise. The quantization noise is amplified along with the information signal, which implies that a $\Delta\Sigma$-based transmitter is not energy efficient. Moreover, although the filtered quantization noise is located outside of the passpand of the information signal, it should be reduced by a sharp bandpass filter to avoid affecting to adjacent bands. For $\Delta\Sigma$-based transmitters, the coefficient of a second-order feedback filter is designed in [6] to reduce the effect of the quantization noise. However, it does not consider the overloading of the 1-bit quantizer. Once an overloading occurs, the $\Delta\Sigma$ modulator becomes unstable and the system may suffer from a burst error. Thus, the overloading should be avoided.

This paper designs a stable 1-bit $\Delta\Sigma$ modulator that minimizes the mean squared error (MSE) of the transmitted signal. Under a condition for the no-overloading of the 1-bit quantizer, the finite impulse response (FIR) filter of the $\Delta\Sigma$ modulator is optimized. Then, the stability of our $\Delta\Sigma$-based transmitter is guaranteed. The 1-bit $\Delta\Sigma$ modulator is designed by using convex optimization problems. Simulations are provided to see the performance of our designed modulator.

Notations: The z transform of a sequence (or a vector) $h = \{h_k\}_{k=0}^{\infty}$ is denoted as $H[\mathrm{z}] = \sum_{k=0}^{\infty} h_k \mathrm{z}^{-k}$. The output sequence y of an linear and time-invariant (LTI) system $H[\mathrm{z}]$ with the input sequence x (i.e. $y = h * x$ where $*$ denotes the convolution) is expressed as $y = H[\mathrm{z}]x$. The l_∞ signal space is defined as the set of all vectors $x = \{x_k\}_{k=0}^{\infty}$ with real components x_k such that $\|x\|_\infty := \sup_k |x_k| < +\infty$. The l_1 norm of an single-input and single-output system $H[\mathrm{z}]$ is defined as $\|H[\mathrm{z}]\| = \sum_{k=0}^{\infty} |h_k|$.

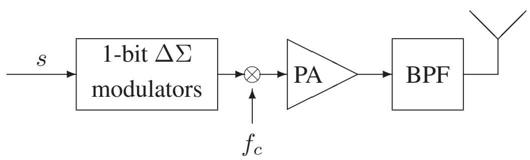

Fig. 1: $\Delta\Sigma$-based transmitter.

2. $\Delta\Sigma$-BASED TRANSMITTERS

Fig. 1 shows a simplified schematic diagram for a $\Delta\Sigma$-based transmitter. The discrete-time complex-valued input s to the $\Delta\Sigma$ modulator is assumed to be bandlimitted in $[-\pi/\text{OSR}, \pi/\text{OSR}]$, where OSR is the oversampling ratio which is a positive integer.

The orthogonal quadrature components of s are independently quantized by two 1-bit $\Delta\Sigma$ modulators. Then, they are upconverted to the carrier frequency f_c. After the RF signal is amplified by the power amplifier (PA), the signal is processed by a band-pass filer (BPF) to reconstruct the information signal as well as to remove the effects of the quantization error.

We only consider the in-phase component of the information symbol, since the orthogonal quadrature components are processed independently in two $\Delta\Sigma$ modulators. We can assume without loss of generality that the input signal has a symmetric magnitude limitation described as

$$\|x\|_\infty \le L_x. \tag{1}$$

Fig. 2 illustrates a $\Delta\Sigma$ modulator, where $Q(\cdot)$ stands for the 1-bit quantizer, $x = \{x_0, x_1, \dots, \}$ and $v = \{v_0, v_1, \dots, \}$ are the input and the output of the $\Delta\Sigma$ modulator, respectively. Let $L > 0$ be the saturation (or equivalently quantization) level. Then, the 1-bit quantizer is defined as

$$Q(u) = \begin{cases} L, & u \ge 0 \\ -L, & u < 0. \end{cases} \tag{2}$$

The overloading occurs if $|u| > 2L$. The round-off error of the 1-bit quantizer is defined as $w = v - u$. If there is no overloading, the round-off error w is bounded such as

$$\|w\|_\infty \le L. \tag{3}$$

Let us denote the filter $R[\text{z}]$ as $R[\text{z}] = \sum_k r_k \text{z}^{-k}$, whose first coefficient r_0 is 1. The round-off error is filtered by the error feedback filter $R[\text{z}] - 1$. The filtered round-off error is fed back to the input to the 1-bit quantizer. Then, the input to the 1-bit quantizer can be expressed as $u = x + (R[\text{z}] - 1)w$. It follows that the output and the quantization error e of the $\Delta\Sigma$ modulator are respectively given by $v = x + R[\text{z}]w$ and $e = v - x = R[\text{z}]w$.

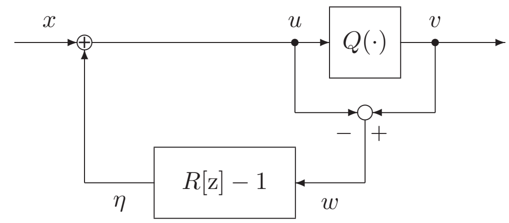

Fig. 2: $\Delta\Sigma$ modulator.

After being processed by 1-bit $\Delta\Sigma$ modulators, the orthogonal quadrature components are modulated to a RF signal, which is amplified by a power amplifier. Since the outputs of 1-bit $\Delta\Sigma$ modulators only take two values, the RF signal has a constant envelop and hence its the peak to average power ratio (PAPR) is a unit. Thus, we can utilize a reasonable and efficient power amplifier. On the other hand, since the quantization noise has to be amplified along with the information signal, the energy efficiency is degraded.

Assuming that the power amplifier is linear, we denote the baseband equivalent discrete-time channel from the output of the $\Delta\Sigma$ modulator to the transmit antenna as $H[\text{z}]$. Then, the transmitted signal y can be expressed in the baseband as

$$y = H[\text{z}]x + \epsilon \tag{4}$$

where ϵ it the filtered quantization noise given by

$$\epsilon = H[\text{z}]e = H[\text{z}]R[\text{z}]w. \tag{5}$$

The filter $R[\text{z}]$ is called a noise shaping filter or a noise transfer function [7]. The filter $H[\text{z}]$ should not change the input signal s except for a delay D, that is, it should be a lowpass filter that satisfies $H[\text{z}]x \approx \text{z}^{-D}x$.

When an overloading occurs, the round-off error w may take a large value. Since the round-off error w is fed back, the input to the 1-bit quantizer can take a large value, which may result in another overloading. Successive overloading may unstabilize the $\Delta\Sigma$ modulator and lead to a burst error at the receiver. Contrary, if there is no overloading, the round-off error w is bounded. Thus, we can summarize that:

Proposition 1. *The 1-bit $\Delta\Sigma$ modulator is bounded-input and bounded-output stable if there is no overloading.*

It is easy to see from the triangle inequality $\|x + \eta\|_\infty \le \|x\|_\infty + \|\eta\|_\infty$ that if

$$L_x + \|R[\text{z}] - 1\|_1 L \le 2L, \tag{6}$$

then no-overloading happens at the 1-bit quantizer. On the other hand, to regulate the power of the quantization noise $e = R[\text{z}]w$, we impose

$$\|R[\text{z}]w\|_\infty = \|R[\text{z}]\|_1 L \le C_e \tag{7}$$

for a positive $c_e \geq L$. We note that $\|R[\mathrm{z}]-1\|_1 = \|R[\mathrm{z}]\|_1 - 1$, since the first coefficient of $R[\mathrm{z}]$ is 1. Then, it follows from (6) and (7) that a sufficient condition for the stability is given by

$$\|R[\mathrm{z}]-1\|_1 \leq \min\left(2 - L_x/L, C_e/L - 1\right). \tag{8}$$

3. DESIGN OF $\Delta\Sigma$ MODULATORS

We would like to design the error feedback filter and find the optimal value for L. However, it is difficult to simultaneously deal with them, since there is the product $\|R[\mathrm{z}]-1\|_1 L$ in the constraint (6).

To obtain a reasonable feedback filter and a reasonable value for L, let us assume that:

Assumption 1. *The quantization error signal w is a white random signal with a zero mean and a variance σ_w^2 and uncorrelated with the input of the uniform quantizer.*

Assumption 1 approximately holds true for uniform quantizers having sufficiently small quantization interval and sufficiently large number of quantization levels. It is often assumed also for $\Delta\Sigma$ modulators, although it is not always satisfied for $\Delta\Sigma$ modulators [8]. Resorting to Assumption 1, let us first design the optimal feedback filter that minimizes the mean squared value for the quantization error ϵ for a given L.

Under Assumption 1, the mean squared error (MSE) of (5) is given by

$$\|H[\mathrm{z}]R[\mathrm{z}]\|_2^2 \sigma_w^2 \tag{9}$$

where the L_2 norm $\|\cdot\|_2$ is defined as

$$\|G[\mathrm{z}]\|_2^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} G^*[e^{j\omega}]G[e^{j\omega}]d\omega \tag{10}$$

with c^* being the complex conjugate of c. The optimal feedback filter that minimizes the MSE is developed in [9]. However no-overloading is not guaranteed.

To minimize the quantization noise, the quantization level L should be set to be its smallest value under the constraint (6). Then, we have

$$L = \frac{L_x}{2 - \|R[\mathrm{z}]-1\|_1}. \tag{11}$$

If the quantization noise is uniformly distributed, the MSE can be expressed as

$$\frac{L_x^2}{3}\frac{\|H[\mathrm{z}]R[\mathrm{z}]\|_2^2}{(2-\|R[\mathrm{z}]-1\|_1)^2}. \tag{12}$$

We cannot minimize the MSE directly. Instead, let us consider the minimization of $\|H[\mathrm{z}]R[\mathrm{z}]\|_2^2$ under the constraint

$$\|R[\mathrm{z}]-1\|_1 \leq \gamma_\eta, \tag{13}$$

which can be formulated as

$$\min_{R[\mathrm{z}]\in RH_\infty, \mu_\epsilon, \gamma_\eta} \mu_\epsilon \tag{14}$$

subject to $R[\infty]=1$, (8) and

$$\|H[\mathrm{z}]R[\mathrm{z}]\|_2^2 < \mu_\epsilon \tag{15}$$

where RH_∞ is the set of stable proper rational functions with real coefficients.

Still, the problem cannot be solved, since the l_1 norm cannot be evaluated easily. To design noise shaping filters, we restrict $R[\mathrm{z}]$ to have a finite impulse response (FIR) and cast the design problem into a convex optimization, which can be solved numerically.

Let $R[\mathrm{z}]$ be an FIR filter of order n, which we denote as $R[\mathrm{z}] = 1+\sum_{k=1}^n r_k \mathrm{z}^{-k}$. The composite system $H[\mathrm{z}]R[\mathrm{z}]$ can be expressed as a state-space realization whose state-space matrices are

$$A = \begin{bmatrix} A_r & B_r C_h \\ \mathbf{0} & A_h \end{bmatrix}, \quad B = \begin{bmatrix} B_r \\ B_h \end{bmatrix} \tag{16}$$

$$C = \begin{bmatrix} C_r & D_r C_h \end{bmatrix}, \quad D = D_h, \tag{17}$$

where (A_h, B_h, C_h, D_h) are state-space matrices of $H[\mathrm{z}]$ and (A_r, B_r, C_r, D_r) are state-space matrices of $R[\mathrm{z}]$ given by

$$A_r = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{bmatrix}, \; B_r = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{18}$$

$$C_r = \begin{bmatrix} r_n, & r_{n-1}, & \cdots & r_1 \end{bmatrix}, D_r = 1. \tag{19}$$

For a given $H[\mathrm{z}]$, the parameters to be optimized are only $(r_1, \ldots, r_n)$.

The inequality (15) can be expressed by linear matrix inequalities (LMIs). It is known that (15) holds true if and only if there exits a positive definite matrix $\mathcal{P}$ that satisfy [10]

$$\begin{bmatrix} \mathcal{P} & \mathcal{P}A & \mathcal{P}B \\ A^T\mathcal{P} & \mathcal{P} & 0 \\ B^T\mathcal{P} & 0 & 1 \end{bmatrix} \succ 0 \tag{20}$$

$$\begin{bmatrix} \mu_\epsilon & C & D \\ C^T & P & 0 \\ D^T & 0 & 1 \end{bmatrix} \succ 0.. \tag{21}$$

On the other hand, the constraint (8) can be written as

$$\sum_{k=1}^n |r_k| \leq \gamma_\eta. \tag{22}$$

Introducing non-negative auxiliary variables $\bar{r}_k \geq 0$ for $k = 1, \ldots, n$ such that $\bar{r}_k = |r_k|$, we can express (22) as in [11]

$$\sum_{k=1}^n \bar{r}_k \leq \gamma_\eta \tag{23}$$

$$-\bar{r}_k \leq r_k \leq \bar{r}_k \quad \text{for} \quad k = 1, \ldots, n \tag{24}$$

$$\bar{r}_k \geq 0 \quad \text{for} \quad k = 1, \ldots, n. \tag{25}$$

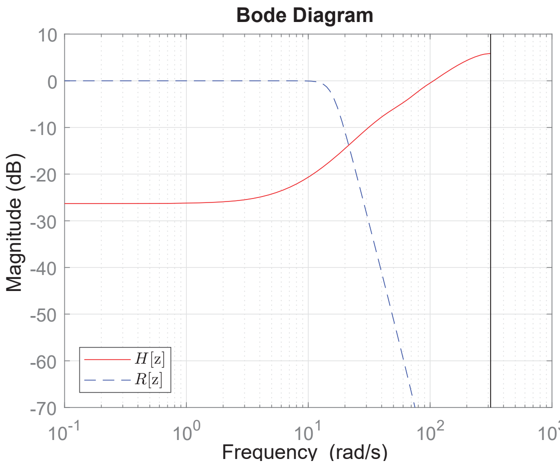

Fig. 3: Frequency responses of $H[\mathrm{z}]$ and $R[\mathrm{z}]$.

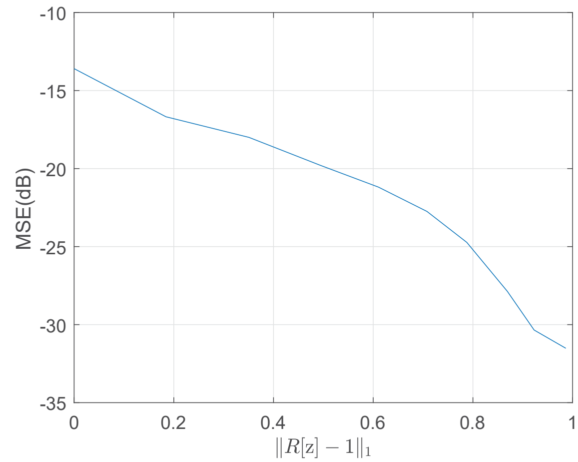

Fig. 4: MSE as a function of $\|R[\mathrm{z}] - 1\|_1$.

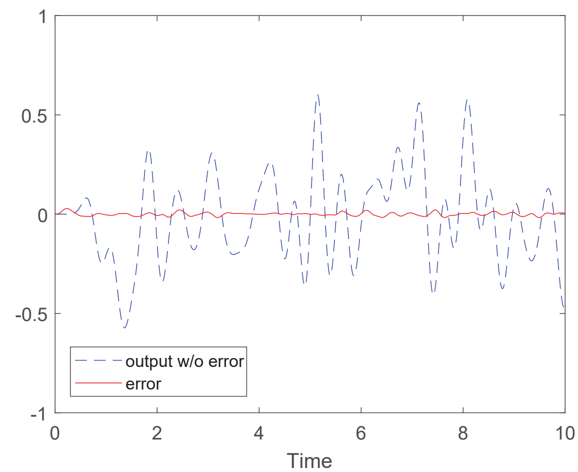

Fig. 5: Output signal $H[\mathrm{z}]x$ and error ϵ in (4).

Since the constraints are convex, the problem is cast into the following convex optimization:

$$\min_{r_1,\ldots,r_n,\bar{r}_1,\ldots,\bar{r}_n,\mu_\epsilon,\gamma_\eta} \mu_\epsilon \tag{26}$$

subject to (20), (21), (23), (24), and (25).

Now, to minimize the MSE (12), we utilize the optimization above. For a fixed upper bound γ_η for $\|R[\mathrm{z}] - 1\|_1$, we can obtain the minimum H_2 norm for $H[\mathrm{z}]R[\mathrm{z}]$ by solving the optimization problem. Changing γ_η from 0 to $\min(2 - L_x/L, C_e/L - 1)$ with a small step size and solving the problem with each value for γ_η, we can numerically find the relationship between $\|R[\mathrm{z}] - 1\|_1$ and $\|H[\mathrm{z}]R[\mathrm{z}]\|_2$, from which we can obtain the minimum MSE and L that achieves the minimum MSE.

4. DESIGN EXAMPLE

For our lowpass filter $H[\mathrm{z}]$, we utilize a 5th-order lowpass Butterworth filter with normalized cutoff frequency 0.05. Fig. 3 depicts the frequency response $H[\mathrm{z}]$ when sampling frequency is 100 Hz.

We generate a lowpass signal by filtering a white noise with an 8th-order lowpass Butterworth filter with normalized cutoff frequency 0.025. The l_∞ norm of the lowpass signal is normalized to be unit, that is, $L_x = 1$ and the length of the FIR filter $R[\mathrm{z}]$ is set to be 15. We set $C_e = 2$ in (7) and then the range of the saturation level L is given by $[1/2, 2]$ from (8). Accordingly, the range of γ_η is $[0, 1]$.

We change the value of γ_η from 0 to 1 with a step size of 0.1. For each value, we solve the optimization problem (26) by CVX [12], a package for specifying and solving convex programs. Fig. 4 shows the MSE as a function of $\|R[\mathrm{z}] - 1\|_1$. We find that $\gamma_\eta = 1$ or, equivalently, $L = 1$ gives the minimum MSE.

We input the lowpass signal x to the designed $\Delta\Sigma$ modulator with $L = 1$ that attains the minimum MSE. Fig. 5 depicts the output signal $H[\mathrm{z}]x$ and the error ϵ in (4). We can conclude that our $\Delta\Sigma$ modulator successively suppress the error due to quantization.

5. CONCLUSION

We have designed an optimal stable 1-bit $\Delta\Sigma$ modulator that minimizes the mean squared quantization error at the output of a $\Delta\Sigma$-based transmitter by solving convex optimization problems. An example is provided to demonstrate our design.

6. REFERENCES

[1] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, “Massive MIMO for next generation wireless systems,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, February 2014.

[2] D. Markert, X. Yu, H. Heimpel, and G. Fischer, “An all-digital, single-bit RF transmitter for massive MIMO,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, no. 3, pp. 696–704, March 2017.

[3] M. M. Ebrahimi, M. Helaoui, and F. M. Ghannouchi, “Delta-sigma-based transmitters: Advantages and disadvantages,” *IEEE Microwave Magazine*, vol. 14, no. 1, pp. 68–78, Jan 2013.

[4] M. Helaoui, S. Hatami, R. Negra, and F. M. Ghannouchi, “A novel architecture of delta-sigma modulator enabling all-digital multiband multistandard RF transmitters design,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 55, no. 11, pp. 1129–1133, Nov 2008.

[5] N. V. Silva, A. S. R. Oliveira, and N. B. Carvalho, “Design and optimization of flexible and coding efficient all-digital RF transmitters,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 1, pp. 625–632, Jan 2013.

[6] U. Gustavsson, T. Eriksson, and C. Fager, “Quantization noise minimization in $\Sigma\Delta$ modulation based RF transmitter architectures,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 57, no. 12, pp. 3082–3091, Dec 2010.

[7] R. Schreier and G. C. Temes, *Understanding Delta-Sigma Data Converters*. Wiley-IEEE Press, 2004.

[8] R. M. Gray, “Quantization noise spectra,” *IEEE Transactions on Information Theory*, vol. 36, no. 6, pp. 1220–1244, Nov 1990.

[9] S. Ohno, T. Shiraki, M. R. Tariq, and M. Nagahara, “Mean squared error analysis of quantizers with error feedback,” *IEEE Transactions on Signal Processing*, vol. 65, no. 22, pp. 5970–5981, Nov 2017.

[10] S. Boyd, L. E. Ghaoul, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Society for Industrial and Applied Mathematics, 1997.

[11] S. Ohno, Y. Ishihara, and M. Nagahara, “Min-max design of feedback quantizers without overloading,” to appear in IEEE Transactions on Circuits and Systems I: Regular Papers.

[12] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming, version 2.1,” http://cvxr.com/cvx, Mar. 2014.