Chapter 1

Higher-order Delta-Sigma Modulators and ABCD Matrix Modeling

In this chapter, a short tutorial on the internal working of the delta-sigma toolbox [1] is provided. A state-space based modeling method is described which allows systematic design and simulation of higher-order $\Delta\Sigma$ modulators.

1.1 Higher-order Delta-Sigma Modulators

Delta-sigma modulators with an order greater than two lead to more aggressive noise shaping and thus better performance. However, these benefits arrive at the cost of more circuitry and reduced input signal range. The generalized delta-sigma modulator is described by a two input loop-filter given by [2]

$$Y(z) = L_0(z)U(z) + L_1(z)V(z) \tag{1.1}$$

where $U(z)$ is the input signal and $V(z)$ is the quantize output fed-back through a ADC. The output of the modulator ($V(z)$) is related to the loop-filter's output ($Y(z)$) by [2]

$$V(z) = Y(z) + E_Q(z) \tag{1.2}$$

where $E_Q(z)$ is the additive linearized model of the quantization noise added by the quantizer. Using the above two equations the modulator's linearized output is given in terms of the inputs as

$$Y(z) = STF(z)U(z) + NTF(z)E_Q(z) \tag{1.3}$$

where we have the noise and signal transfer functions given by [2]

$$NTF(z) = \frac{1}{1 - L_1(z)}$$

$$STF(z) = \frac{L_0(z)}{1 - L_1(z)} \tag{1.4}$$

On the other hand, we can express the loop transfer functions in terms of NTF and STF as

$$\begin{aligned} L_0(z) &= \frac{STF(z)}{NTF(z)} \\ L_1(z) &= 1 - \frac{1}{NTF(z)} \end{aligned} \tag{1.5}$$

The above described generalized modulator structure is shown in the block diagram in Figure 1.1.

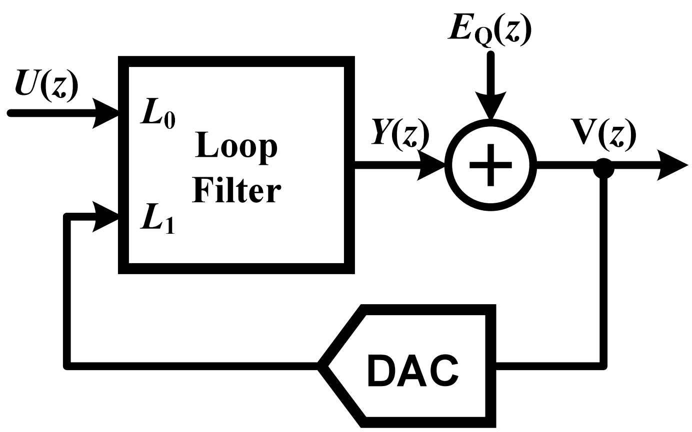

Figure 1.1: A generalized structure of a delta-sigma modulator.

From Equation 1.5, we can deduce that for the NTF to be low in the signal band ($f = 0$ to f_B, where $f_B = \frac{f_s}{2 \cdot OSR}$) $L_1(z)$ must be large in the signal band with a low-pass response. Also $L_0(z)$ should be large in the signal band so as to maintain the STF response close to unity. This implies that both L_0 and L_1 must have poles in the same vicinity. Since L_0 and L_1 share the same circuitry, these poles are usually co-incident and form the zero of the NTF. However, the zeros of L_0 and L_1 are generally distinct. In the simplest case, for a L^{th} order modulator, where the $NTF(z) = \left(1 - z^{-1}\right)^L$ differentiates the quantization noise M-times and $STF(z) = z^{-k}$, $k < L$. This leads to the expressions

$$\begin{aligned} L_0(z) &= z^{-k}\left(1 - z^{-1}\right)^{-L} \\ L_1(z) &= 1 - \left(1 - z^{-1}\right)^{-L} \end{aligned} \tag{1.6}$$

where the poles of both L_0 and L_1 are located on the unit circle , at $z = 1$. On the other hand, for L_0, the $(L-k)$ zeros are located at $z = 0$ while the remaining k zeros are located at infinity [2]. The zeros of L_1 are given by the roots of the equation $\left(1 - z^{-1}\right)^L = 1$ and are expressed by the expression

$$z_l = \left(1 - e^{-\frac{j2\pi l}{L}}\right) = \frac{(1 + j\cot(\pi l/L))}{2}, \;\; l = 1, 2, ..., L-1 \tag{1.7}$$

For this modulator and with $OSR \gg 1$, the in-band noise power is given by

$$\sigma_n^2 = \frac{\sigma_q^2 \pi^{2L}}{(2L+1)\, OSR^{2L+1}} \tag{1.8}$$

and the maximum signal-to-quantization noise ratio (SQ NR) is expressed as [2]

$$SQNR = 6.02N + 1.76 + (20L+10)\log_{10}(OSR) - 10\log_{10}\left(\frac{\pi^{2L}}{2L+1}\right) \tag{1.9}$$

From the above equation, it can be deduced that for a higher oversampling ratio (OSR), the SQNR increases by $\left(L + \frac{1}{2}\right)$ bits per doubling in OSR. Thus increasing the order of the modulator has a direct impact on the achievable resolution.

1.2 NTF Pole and Zero Optimization

NTF pole and zero optimization is a an important technique for synthesis of high-resolution, higher-order modulators. In the simple example seen in Section 1.1 the NTF was of the form $\left(1 - z^{-1}\right)^L$ and all the zeros were located at $z = 1$ and the poles were at $z = 0$. Now if we intelligently spread these zeros around in the signal band and also move the poles to surround the zeros while being within the unit circle, significant improvements in SQNR can be achieved. This NTF pole and zero optimization is done using the delta-sigma toolbox [1] in Matlab and is its algorithm is detailed in [2]. Figure 1.2 compares the spectra of a single-bit third-order DSM with $OSR = 64$ with and without NTF (zero and pole) optimization. The locations of the NTF zeros as the result of the SQNR optimization are shown in Figure 1.3.

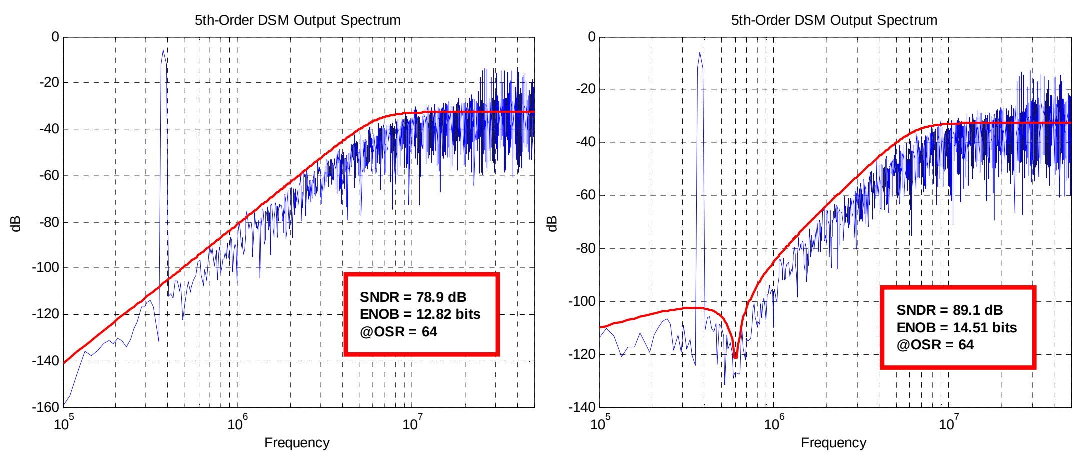

Figure 1.2: Example of NTF zero and pole optimization. Here a third-order single-bit delta-sigma modulator is synthesized for $OSR = 64$.

Here, we can observe that the SQNR is increased by roughly 10 dB resulting in 1.7 bits increase in resolution. Further, we observe that the noise floor in the later case is raised and the NTF notch in the signal band is flattened out. This relaxes the gain requirements on the op-amps in the loop-filter.

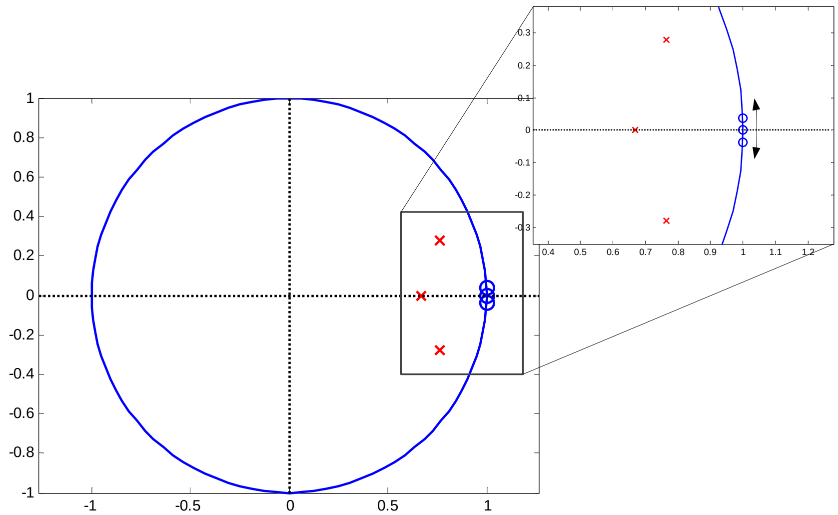

Figure 1.3: NTF zero spreading in the signal band with SQNR optimization.

1.3 Loop-filter Architectures

There are many generalized structures to realize the higher-order delta-sigma modulators. The first loop-filter architecture, shown in Figure 1.4, is called Cascade of Integrators with Distributed FeedBack (CIFB) [2]. This topology comprises of a cascade of L delaying integrators. The feedback and signal inputs are fed into each of the integrators with weight factor coefficients a_i and b_i ($a_i, b_i > 0$). For now, assume that we have $g_i = 0$ and $c_i = 1$. The transfer function for the input loop-filter L_0 is given by [2]

$$\begin{aligned} L_0(z) &= \sum_{i=1}^{N+1} \frac{b_i}{(z-1)^{N+1-i}} \\ &= \frac{b_1 + b_2(z-1) + \cdots + b_{N+1}(z-1)^N}{(z-1)^N} \end{aligned} \tag{1.10}$$

and the loop-filter seen by the feedback is given as [2]

$$\begin{aligned} L_1(z) &= \sum_{i=1}^{N} \frac{-a_i}{(z-1)^{N+1-i}} \\ &= -\frac{a_1 + a_2(z-1) + \cdots + a_N(z-1)^{N-1}}{(z-1)^N} \end{aligned} \tag{1.11}$$

Thus the noise transfer function NTF of the modulator is given by [2]

$$NTF(z) = \frac{1}{1 - L_1(z)} = \frac{(z-1)^N}{D(z)} \tag{1.12}$$

where the denominator is expressed as

$$D(z) = a_1 + a_2 (z-1) + \cdots + a_N (z-1)^{N-1} + (z-1)^N \tag{1.13}$$

From the above two equations, we can observe that the zeros of the NTF are located at $z = 1$ which is at DC. The locations of the poles is controlled by the coefficients a_i. A;so, the signal transfer function is given as

$$STF(z) = \frac{L_0(z)}{1 - L_1(z)} = \frac{b_1 + b_2 (z-1) + \cdots + b_{N+1} (z-1)^N}{D(z)} \tag{1.14}$$

Here we see that the poles of the STF are same as that of the NTF and are determined by a_i. On the other hand, the zeros of the STF are determined by b_i [2].

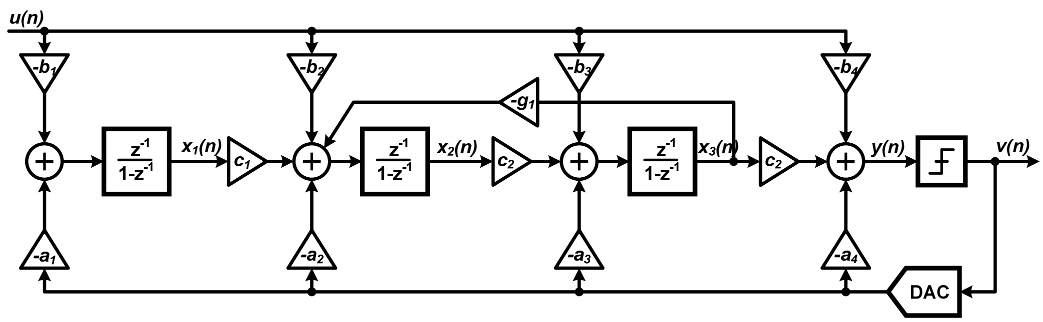

Figure 1.4: CIFB (Cascade of Integrators with Distributed FeedBack) modulator topology.

An important case is when $b_i = a_i$ for $\forall i \leq N$ and $b_{N+1} = 1$, then the STF is exactly equal to 1 and the modulator's output is equal to

$$V(z) = U(z) + NTF(z)E_Q(z) \tag{1.15}$$

For this condition, the modulator input signal $u[n]$ is not processed by any of the integrators. The loop filter only acts on the quantization noise $e(n)$. Due to this the integrator swings are reduced and the integrator non-linearities do not introduce distortion into the signal path. The amount of quantization noise processed by the loop-filter is further reduced when multi-bit quantizers are employed [2].

So far the NTF zeros have been designed to be located at DC ($z = 1$). In Section 1.2, it was shown that by distributing these zeros appropriately in the signal band, significant improvements in SQNR can be achieved. Now, in Figure 1.4, consider g_i to be non-zero. Now, we observe that the middle two integrators, together form a second-order resonator with the locations of the zeroes being determined by g_1. The transfer function of the first resonator in the figure is given by [2]

$$R_1(z) = -\frac{a_1 z + a_2(z-1)}{z^2 - 2z + (1+g_1)} \tag{1.16}$$

here the zeroes are located at $z = 1 \pm j\sqrt{g_1}$. Even though the resonators are individually unstable, when embedded in a well designed feedback system, they do not exhibit local oscillations. In Figure 1.4, the coefficients c_i are obtained as a result of dynamic range scaling (DRS), discussed later in

Section 1.4.2. A systematic loop-filter design method with NTF zero optimization is detailed in [2] and is implemented in the delta-sigma toolbox [1].

Instead of using a feedback structure to realize the loop-filters L_0 and L_1 given by Equations 1.10 and 1.11 respectively, a feed-forward structure can also be employed. Figure 1.5 shows a third-order modulator with delaying integrators and feed-forward branches. this topology is called **C**hain of **I**ntegrators with **F**eed-**F**orward summation (CIFF) .

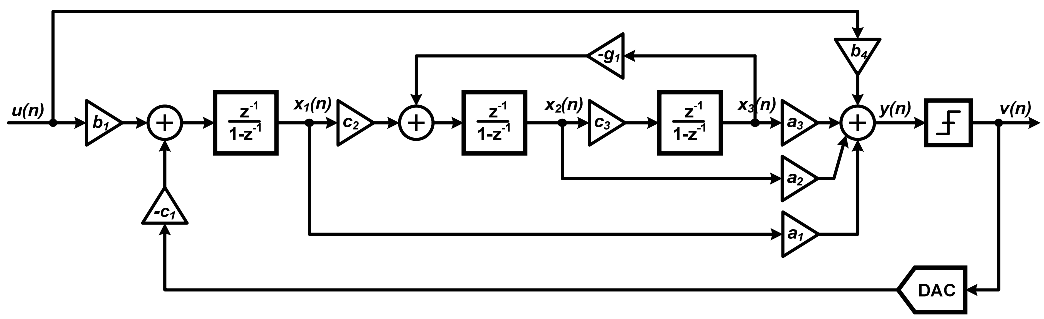

Figure 1.5: CIFF (Cascade of Integrators with distributed FeedForward) modulator topology.

We can observe that the feedback filter transfer function is given by [2]

$$L_1(z) = -a_1 I(z) - a_2 I^2(z) - \cdots - a_N I^N(z) \tag{1.17}$$

where $I(z) = \frac{z^{-1}}{1-z^{-1}}$ is the transfer function of the delaying integrator. Similarly, the input transfer function is given by [2]

$$L_0(z) = b_1 \left(\sum_{i=1}^{N} a_i I^i(z) \right) + b_2 \left(\sum_{i=2}^{N} a_i I^i(z) \right) + \cdots + b_{N+1} \tag{1.18}$$

For this structure when $b_i = 0,\ i = 2, \ldots, N$ and $b_1 = b_{N+1} = 1$, we have $L_0(z) = 1 - L_1(z)$ which leads to $STF(z) = 1$. Again, for this condition, the input to the loop-filter satisfies

$$U(z) - V(z) = -NTF(z)E_Q(z) \tag{1.19}$$

which implies that the loop-filter processes only the quantization noise and not the input signal. Thus, this loop-filter has the property of low distortion and relaxed performance requirement for the op-amps. The branches g_i are added for NTF zero optimization and c_i introduced to accommodate dynamic range scaling.

1.4 Synthesis procedure for ΔΣ modulators

The Schreier's ΔΣ Toolbox in Matlab is widely used to rapidly synthesize and simulate delta-sigma modulator topologies. The synthesis process yields the loop-filter coefficients for the modular topology designed for a given set of specifications. These resulting filter co-efficients are mapped to a switched-capacitor filter for discrete-time implementation, and to either op-amp-R or g_m-C filter for the continuous-time implementation of the modulator. The toolbox internally employs constructs and techniques from linear systems theory to describe the loop-filter topologies. In this section a short background on the working of the toolbox is provided. Further this toolbox is extended to synthesize and simulate generalized higher-order $\Delta\Sigma$ modulators.

1.4.1 The ABCD Matrix

The delta-sigma toolbox internally uses ABCD matrix to represent the linear part of the modulator, which are the loop-filters L_0 and L_1, as illustrated in 1.6. The ABCD matrix representation of the loop-filter is indispensable for linear operations like dynamic range scaling, automated design mapping and for rapid discrete-time simulation of the modulator architectures. The ABCD matrix is a combination of four sub-matrices which describe the dynamics of any discrete-time linear system. The state-space equations for the DSM loop filter are described as

$$\begin{aligned} \mathbf{x}[n+1] &= A\mathbf{x}[n] + B\begin{bmatrix} u[n] \\ v[n] \end{bmatrix} \\ \mathbf{y}[n] &= C\mathbf{x}[n] + D\begin{bmatrix} u[n] \\ v[n] \end{bmatrix} \end{aligned} \tag{1.20}$$

where $\mathbf{x}(n) \in R^{M\times 1}$ is the state vector at time n for an M^{th}-order modulator. The matrix $A \in R^{M\times M}$ defines the interconnections withing the loop filter. The matrix $B \in R^{M\times 2}$ describes how the modulator input $u[n]$ and the feedback DAC output $v[n]$ are applied to the loop filter $H(z)$. The matrices $C \in R^{1\times M}$ and $D \in R^{1\times 2}$ describe the computation of the output $y[n]$ from the states $\mathbf{x}[n]$ and the loop filter inputs $\begin{pmatrix} u[n] & v[n] \end{pmatrix}^T$ [2].

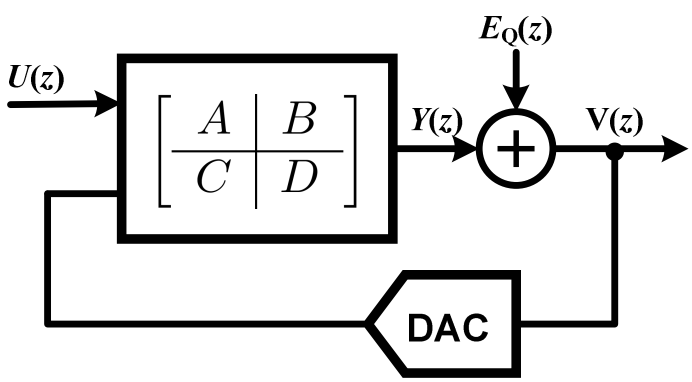

Figure 1.6: The ABCD Matrix representation of the loop-filter in a delta-sigma modulator.

The loop transfer functions are obtained from the ABCD matrix as

$$\begin{bmatrix} L_0(z) \\ L_1(z) \end{bmatrix} = C\left(z^{-1}I - A\right)^{-1} B + D \tag{1.21}$$

The delta-sigma toolbox evaluates the equivalent closed-loop ABCD matrix, $ABCD_{cl}$ as

$$\begin{aligned} A_{cl} &= A + k_q B_2 C \\ B_{cl} &= \begin{bmatrix} B_1 + k_q B_2 D_1 & B_2 \end{bmatrix} \end{aligned}$$

$$
\begin{aligned}
C_{cl} &= k_q C \\
D_{cl} &= \begin{bmatrix} k_q D_1 & 1 \end{bmatrix}
\end{aligned}
\tag{1.22}
$$

where $B_{1,2}$ are columns of $B = \begin{bmatrix} B_1 & B_2 \end{bmatrix}$ and D_1 is the first column element of $D = \begin{bmatrix} D_1 & D_2 \end{bmatrix}$. Here, it is assumed that the quantizer delay ($z^{-1/2}$) has been absorbed into the last integrator delay of $z^{-1/2}$ and thus the quantizer only introduces a gain of k_q and has zero delay. Consequently, from the $ABCD_{cl}$ matrices the STF and NTF are evaluated as

$$
\begin{bmatrix} STF(z) \\ NTF(z) \end{bmatrix} = C_{cl} \left(z^{-1}I - A_{cl}\right)^{-1} B_{cl} + D_{cl} \tag{1.23}
$$

1.4.2 Dynamic Range Scaling

Dynamic range scaling (DRS) is an important step when designing practical $\Delta\Sigma$ modulators. In DRS, the ABCD matrix of the loop-filter is scaled so that the individual state maxima are bounded by a specified limit x_{lim}. The value of x_{lim} is selected such that the op-amp outputs lie within the $x_{lim} \cdot V_{DD}$ range and linear operation of the loop-filter is assured. This value is usually selected to be around $\frac{1}{3}$ to $\frac{1}{2}$ depending upon the op-amp design. The maximum stable amplitude (u_{max}) is also obtained as a result of this scaling process. In the range scaling process, first the ratios $r_i = \frac{x_{max,i}}{x_{lim}}$ of the state maxima $x_{max,i}$ to x_{lim} are estimated through transient simulations. Then, the diagonal scaling matrix S is formed with the inverse these ratios and is given as [1]

$$
S = \begin{bmatrix} \frac{1}{r_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{r_2} & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & \frac{1}{r_M} \end{bmatrix} \tag{1.24}
$$

Then S is applied on the state vector to obtain the scaled state vector $\mathbf{x_s} = S\mathbf{x}$. This ensures that all the states are bounded within x_{lim}. The resulting ABCD matrix after range scaling is given by

$$
ABCD_s = \left[\begin{array}{c|c} SAS^{-1} & SB \\ \hline CS^{-1} & D \end{array}\right] \tag{1.25}
$$

The above described range scaling process is illustrated in Figure 1.7 where a single state in the loop-filter is range scaled by r.

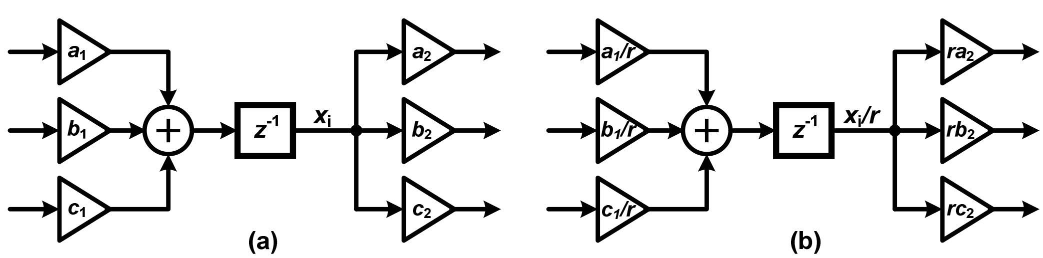

Figure 1.7: Dynamic range scaling of a loop-filter state from x_i to x_i/r.

1.4.3 Mapping to a Loop-filter Architecture

To summarize the synthesis procedure developed in the chapter, the process of designing a generalized higher-order $\Delta\Sigma$ modulator can be summarized as follows:

1. Using the delta-sigma toolbox, synthesize a NTF-zero optimized, $\Delta\Sigma$ modulator for a given order with an oversampling ratio equal to OSR.

2. Apply dynamic range scaling on the $ABCD$ matrix of the synthesized loop-filter to restrict the loop-filter states to x_{lim}.

3. Map the scaled $ABCD$ matrix to either CIFB (CRFB) or CIFF (CRFF) modulator topology.

4. Using simulations, estimate the effective gain of the quantizer (k_q) employed in the modulator as $k_q = \frac{E[|y|]}{E[y^2]}$.

5. Evaluate the NTF and STF for the synthesized $\Delta\Sigma$ modulator using the estimated k_q value (close to 1 for higher-order modulators).

Bibliography

[1] R. Schreier. The Delta-Sigma Toolbox.
Internet:http://www.mathworks.com/matlabcentral/fileexchange/19.

[2] R. Schreier and G.C. Temes. *Understanding Delta-Sigma Data Converters*. IEEE press Piscataway, NJ, 2005.