System Identification of Sigma Delta Modulators for Performance Verification

Ronan Farrell

Department of Electronic Engineering
National University of Ireland, Maynooth
IRELAND
E-mail: ronan.farrell@may.ie

Abstract – **A technique is presented that allows for the easy and rapid extraction of the parameters of a sigma-delta modulator system. While an interesting problem in the area of mixed signal system identification it also has applications in the silicon device production environment.**

Keywords – **sigma delta modulators, test, system identification**

I INTRODUCTION

For very high-resolution performance at audio frequencies, for example CD quality music applications, sigma-delta modulator based analog-digital data converters are generally the optimum choice [1]. Sigma-delta based converters sample the incoming data at much higher frequencies than needed. Through modulation, system imperfections are converted to noise and shifted to higher frequencies. This noise is then removed through additional digital filtering. This approach enables resolutions in excess of 16 bits to be obtained.

However this technique leads to difficulty in proving that a sigma-delta modulator system has been manufactured correctly and that it satisfies the required performance specifications. The standard techniques for testing analog-to-digital converters (ADCs) are not appropriate for use with high performance sigma-delta modulator.

This paper will be divided into several sections, the first gives a brief description of sigma-delta modulators and standard ADC test techniques. The second section introduces a new technique that allows for rapid identification of the system's performance. The next section shows the results of simulations that illustrates the performance of this approach. The final sections will discuss this work, future developments and conclusions.

II SIGMA-DELTA MODULATORS

Sigma-delta modulators have become very popular since the 1980's and are now generally accepted as the optimum data converter for high performance low to audio frequency applications. This is due to their feedback architecture that allows the individual components of the system to be more tolerant of manufacturing imperfections [2]. In recent times, sigma-delta modulators have been applied to higher frequency applications, including an increasing number of video-rate applications [3]. Their range of application is limited only by the fact that they must be used in conjunction with significant oversampling which places tight restraints on their maximum sampling frequency.

Sigma-delta modulators were first suggested by Cutler in 1960 [4], with the first published description of their properties was by Inose and Yasuda [5] three years later. Sigma-delta modulators are based upon the principle of using feedback to improve the effective resolution of a coarse quantiser, commonly a 1-bit quantiser. If this feedback is used in conjunction with discrete-time integrators then the mean of the output of the quantiser will oscillate about the mean of the input, these oscillations decreasing in size over a number of iterations. Each output is a coarse quantisation of the input, depending on the resolution of the quantiser. This combined with the technique of oversampling results in a high-resolution composite output.

This approach also tends to produce a high-pass noise transfer function, shifting the noise away from the low frequencies of interest to higher frequencies where the noise can be filtered away. The configuration of the integrators and the feedback loops from the output of the quantiser determine the noise-shaping effect of the modulator. It is common to design the system to obtain low-pass noise-shaping but it is possible to obtain band-pass noise-shaping for use in high frequency applications [6].

The noise-shaping of the sigma-delta modulator can be enhanced by using a more complex architecture, increasing the number of integrators, and adding additional feedback loops from the quantiser to the inputs of these new integrators.

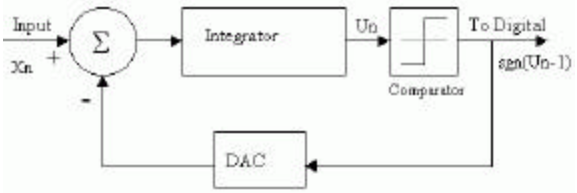

Figure 1: First order sigma-delta modulator

The most basic sigma-delta modulator is the first-order single loop modulator, shown in Figure 1. It consists of a 1-bit quantiser, a discrete-time integrator (which is modelled by a unit delay and a feedback loop), and a feedback loop from the output of the quantiser to the input via a DAC. A one-bit DAC is very simple to implement, often no more than a switch that changes the feedback path from being connected to a high or a low voltage. The behaviour of the first-order sigma-delta modulator is given by a single difference equation:

$$u_n = u_{n-1} + x_n - \mathrm{sgn}(u_{n-1}) \qquad (1)$$

where

u_n = value stored on the integrator

x_n = the input

$$\mathrm{sgn}(y) = \begin{cases} -1 & when\ y < 0 \\ +1 & when\ y \geq 0 \end{cases}$$

However this is an ideal representation of the first-order sigma-delta modulator. Practical implementations will introduce errors which need to be modeled. These errors are common to all sigma-delta architectures, though here they will be described only in terms of the first order system. This is for two reasons

- The first order system is the simplest yet retains all the major sources of implementation error, thus clarity of presentation is preserved
- In higher order systems, the most dominant errors are those that occur closest to the input. All other errors are reduced by the loopgain of the system, and can often be ignored when compared to those of the first integrator stage, thus the first order system is a good working model.

There are four main errors that can be introduced when implementing a sigma-delta modulator.

1. Scalar error introduced between the input and the integrator input.
2. Scalar error introduced between the expected ideal DAC output and the integrator input.
3. Offset error introduced between the expected ideal DAC output and the integrator input.
4. Non-infinite gain on the operational amplifier which is used to form the discrete-time integrator. [7]

By examining a typical implementation of a first order system, the source of these errors will become more apparent. The more experienced engineers will appreciate that these systems are implemented generally as differential-mode circuits rather than as single-ended circuits for noise purposes but the dominant errors are the same.

A typical implementation of a first-order sigma-delta modulator resolves about the discrete-time integrator. The most common implementation is that of a switched-capacitor implementation. (figure 2).

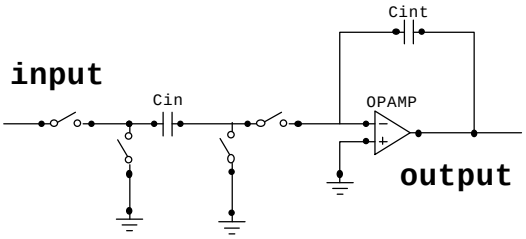

Figure 2: A switched-capacitor integrator.

This circuit implements a unit-delayed discrete time integrator with the scaling factor given as the ratio of the two capacitors. For an ideal unit-gain discrete-time integrator, the capacitors should be equal. The transfer function of this circuit is given below for the case of infinite amplifier gain.

$$H(z) = \left(\frac{C_{in}}{C_{\mathrm{int}}}\right)\left(\frac{z^{-1}}{1-z^{-1}}\right) \qquad (2)$$

However when the gain of the amplifier is not infinite, the charge stored on the capacitor about the integrator is not maintained correctly and on each integration-step the charge decays slightly due to incomplete settling [7]. This modifies the transfer function of (2) to

$$V_{out} = \left(\left(\frac{C_{in}}{C_{\mathrm{int}}}\right)V_{in} + V_{stored}\right)\frac{1}{1-1/A} \qquad (3)$$

where

A = amplifier gain.

Examining (3) more closely, it is possible to see that the integrator not only has a nonlinear scaling of its own stored value due to the amplifier gain, but this also introduces a small scaling factor on the input value. The non-linear effect of the amplifier gain is actually more significant than this expression would suggest due to its effect in the overall non-linear system. Figure 3 shows the resulting distortion in what should be an ideally linear relationship between

the input and output. This is often called the devils staircase and is common in non-linear dynamics [7].

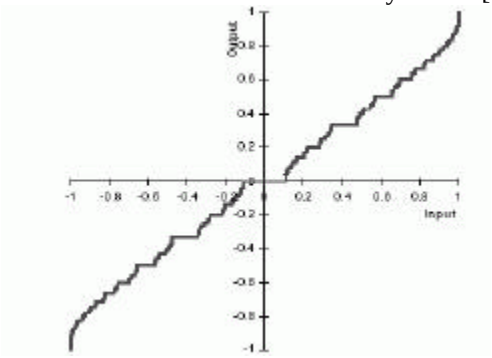

Figure 3: Non-linearities in the output transfer function due to finite amplifier gain.

A more significant scaling effect will arise due to mismatch in the capacitor values. In typical chip-manufacturing processes available today, it is difficult to match any two capacitors better than 1%.

One of the reasons that switched-capacitor implementations are more common is that the addition function before the integrator (Figure 1) can be easily implemented as shown (Figure 4)

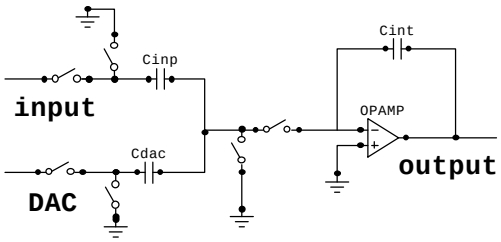

Figure 4: Integrator combined with addition function

In this case, the scaling effect is more critical as it changes the ratio of the input to the feedback value from the DAC. This is highly critical as the output value actually represents the value of the input as a fraction of the DAC feedback value. Thus a mismatch between C_{DAC} and C_{in} would result in a system gain error.

Another source of error that needs to be considered is that this value produced by the DAC may be incorrect. It may not be the expected full range signal due to a gain error, and it may be offset from the expected range. Any unexpected scaling from the DAC will result in a system gain error, and an offset will offset the input by the same amount in the opposite direction.

These errors when combined can be included in the system equation (1):

$$u_n = Gu_{n-1} + K_A x_n - K_B\left(\mathrm{sgn}(u_{n-1}) + \Delta_{DAC}\right) \quad (4)$$

where

G represents the leakage in the stored charge 1/(1-1/A)

III PERFORMANCE VERIFICATION

In a production environment it is important to be able to ascertain whether the system is operating normally. Typically the only values available for measurement are the input and the output. The input is typically ramped across all possible output codes and the output is then compared with the input. If it satisfies this test, and others, then it is deemed to be a good device.

However this approach impedes rapid testing. Typically a high-resolution sigma-delta modulator converter will take 128 samples for each output and will require three times that number before a typical output filter ($Sinc^3$ is optimum) will produce a valid output. If each code of a 16-bit converter is to be test, 8 million clock cycles would be required. This is inelegant and is also an unrealistic expectation in a production environment. In practice only a small set of sample codes are examined. However this does run a risk of missing an error.

This paper proposes a new test that will quickly identify those devices that do not conform to the required specification prior to more intensive testing.

IV SYSTEM IDENTIFICATION

This paper addresses the case where the failure is due to non-linearities arising from finite amplifier gain or offsets and scaling in one or more for the input paths to the integrators as these represent the majority of catastrophic failures. These errors can result in system gain and offset errors, instability and non-idealities as shown in Figure 3. If instead of monitoring system output performance, the system parameters can be characterised, limits on these parameters can be used to provide initial indication of overall system performance.

It is possible by the monitoring of the input value, the digital output from the *sgn* function and one additional internal variable, the output of the integrator, to fully characterise the system by taking advantage of the behaviour of the output bitstream and how this corresponds to internal behaviour. It is shown that instead of requiring hundreds of cycles to obtain a single output as in the traditional approach, this approach will provide complete system characterisation in fewer than one hundred clock cycles.

Sigma-delta modulators, of all orders, have pseudo-random output bitstreams from the *sgn* function. The mean of this bitstream equals the mean of the input. However these cycles for inputs within the centre 30% of the input range [8] have short cycles, rapidly interchanging between high and low outputs (bits 1 or 0). The approach being presented in this paper

uses this behaviour to help simplify the system and thus extract the system parameters.

a) Calculating K_A and K_B

The integrator output can be expressed as a summation of previous inputs to the integrator. over a number of iterations, it is possible to say that

$$u_{n+k} = u_n + \sum_{i=0}^{k-1} \left(x_{n+i} - \mathrm{sgn}(u_{n-1+1}) \right) \tag{5}$$

Or with the imperfections

$$u_{n+k} = G^k u_n + \sum_{i=0}^{k-1} \left(K_A x_{n+i} \right) - \sum_{i=0}^{k-1} K_B \left(\mathrm{sgn}(u_{n-1+1}) + \Delta \right) \tag{6}$$

For simplicity, rename $sgn(u_n)$ Q_n.

It can be noted that u_{n+1} can be expressed in term of known components *x(n)* and *Q(n)* and thus (6) can be expressed more fully.

$$u_{n+k} = G^k u_n + \left(\begin{array}{l} + \sum_{i=0}^{k-1} G^i K_A x_{n+k-i} \\ + \sum_{i=0}^{k-1} G^i K_B \Delta \\ + \sum_{i=0}^{k-1} G^i K_B Q_{n+k-i} \end{array} \right) \tag{7}$$

In practice the gain of any system will be in excess of several hundred if not in excess of a thousand. Thus it can be easily seen that the deviation of G from 1 will be small and for higher powers of G will be insignificant. For this reason, in the following section, the value of G^n will be approximated as being equal to one.

Presume that it is possible to identify a set of outputs Q_n such that the sum of these outputs equals zero (*SQ=0*). In such a case the summation of these values in (7) equals zero. It is shown in [9] that such a cycle of length 6 exists and is frequently found in the output bitstream for a centre-of-range input value (a value in the centre 30% of the input range).

$$\begin{aligned} u_{n+k} &= u_n + \sum_{i=0}^{k-1} \left(K_A x_{n+k-i} + K_B \Delta \right) \\ &= u_n + M \end{aligned} \tag{8}$$

M is measurable as *u* and *x* are known. However more data is needed before any progress can be made in resolving M into its components.

Allow dU to be the change in u over the monitored period and the summation of the values of input (x) over the same period be ΣX. If the input is constant over this period, then multiple measurements will reduce the effect of any noise in any measurements. The measurement of a change in u over a number of cycles is more dependent on noise but if the change is large, the effect of noise is minimised. Measure dU and ΣX for two different values of input, giving dU_1, dU_2, ΣX_1, ΣX_2.

Values for M can be found as being the equal to the value of dU for a given period, or preferably the mean value of dU over a number of monitoring periods. With these values, the following expressions can be developed:

$$K_B \Delta = \frac{M_1 \Sigma X_2 - M_2 \Sigma X_1}{\Sigma X_1 - \Sigma X_2} \tag{9}$$

$$K_A = \frac{M_1 - 6 K_B \Delta}{\Sigma X_1}$$

Thus it is possible to isolate K_A and $K_B\Delta$.

Identifying the value of K_B is a simple task. Another section of the bitstream is selected, except in this case, let the summation of the quantiser outputs Q_n sum to two (*SQ=2*). If the assumption is made that U_n takes an approximately equal distribution of values for when *SQ=0* and when *SQ=2*, then the change in the measured change in U over an equal sized period is solely due to the effect of a different value of quantiser. Then with the assumption of |Q|=1, we can define K_B as

$$K_B = \frac{M_{SUM=2} - M_{SUM=0}}{2} \tag{10}$$

It can be shown that for a given input, the distribution pattern for values of U are equally distributed, given a large enough set of samples. This is not a stringent requirement and simulations indicate that within 100 cycles sufficient equality is achieved.

In the coming diagrams, the results for this approach are presented. In any practical application there will be a limit to the measurement resolution. The graphs will show that by the approach taken, the results will degrade gracefully with reductions in measurement resolution. The graphs have been calculated using a sample modulator with the following parameters.

```
KA = 0.7123289
KB = 1.31123289
Δ  = 0.0447  (offset)
X1 = 0.12117267;
X2 = 0.31234567;
A  = 100,000
```

There is no particular requirement on the two input values used except that better performance results if the two are not close. However the values should be chosen in light of the requirement that the required bitstream segments must be present.

Figure 3 shows the error in the extracted parameters as the errors are so small as be indistinguishable when viewed with the original values.

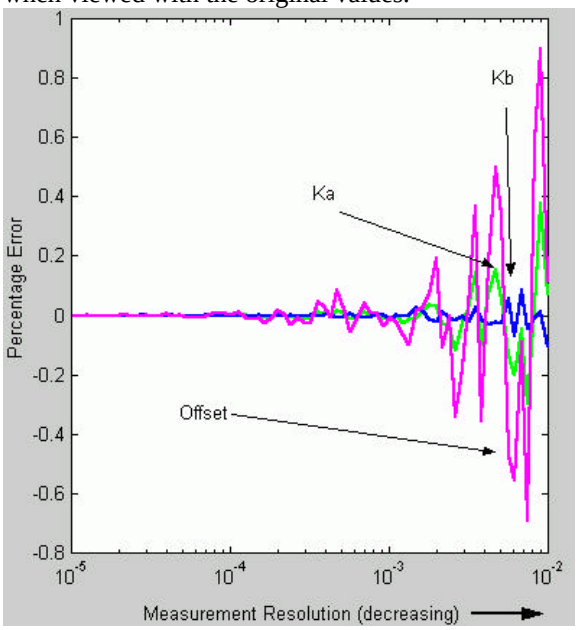

Figure 5: Percentage Errors in extracted parameters, averaged over 256 cycles.

As can be seen, the errors are small, less than 1% even with a measurement resolution no better than 1% (compared to a maximum input value of 1.0) and with noise at the 1% level. The exact performance would be improved in the absence of noise, however this is not easily achieved. Better performance can be obtained by averaging over a larger number of cycles. As the monitoring period is extended, more eligible bitstream segments are generated. Noise and measurement resolution are not a significant factor in extracting these parameters.

b) Determining the Presence of Finite Gain

The final important characteristic to identify is the gain value of the amplifier. Though ignored earlier, it is important to be able to determine whether the value of the amplifier gain is sufficiently large to be acceptable for the performance required. The approach used for this depends on the results already previously obtained for K_A, K_B and Δ. We propose that the output bitstream should be matched with a short series of outputs (length N) from a mathematical model of an ideal sigma-delta modulator using the discovered values of K_A, K_B and Δ but with an ideal amplifier. The mathematical model is seeded with the measured value of the integrator output. Given the condition that the output bitstream is identical (the quantiser feedback contributions are equal), any difference in the output of the integrator after N cycles will be due to either errors in the extracted parameters or due to finite amplifier gain. Assuming parameter contributions are small, the primary source of error will be finite gain. Consider the following expressions for the integrator output.

$$u_{n+k} = G^k u_n + \sum_{i=0}^{k-1} \begin{pmatrix} G^i K_A x_{n+k-i} \\ + G^i K_B \Delta \\ + G^i K_B Q_{n+k-i} \end{pmatrix} \quad (11)$$

$$u_{n+k(ideal)} = u_n + \sum_{i=0}^{k-1} \begin{pmatrix} K_A x_{n+k-i} \\ + K_B \Delta \\ + K_B Q_{n+k-i} \end{pmatrix} \quad (12)$$

Assuming that we have the same period of time *k*, and quantiser output pattern, then the difference in the two is given as

$$\Delta U = \left(1-G^k\right)u_n + \sum_{i=0}^{k-1} \begin{pmatrix} \left(1-G^i\right)K_A x_{n+k-i} \\ + \left(1-G^i\right)K_B \Delta \\ + \left(1-G^i\right)K_B Q_{n+k-i} \end{pmatrix}$$

$$\cong \frac{k}{A}u_n + \frac{1}{A}\sum_{i=0}^{k-1} \begin{pmatrix} iK_A x_{n+k-i} \\ iK_B \Delta \\ iK_B Q_{n+k-i} \end{pmatrix} \quad (13)$$

as

$$G = 1 - \frac{1}{A}$$

$$1 - G^k = \frac{k}{A}$$

Thus finite gain effects will build up with increasing segment length. The errors will be particularly large if the initial stage of the comparison sequence has large values for the integrator output. However if midrange values are used that experience rapid changes in the output bitstream, the results may be masked by negative and positive valued errors destructively cancelling. Simulation shows that a relatively short period can be more effective (Figure 6). This graph shows the error for a range of gain values and input values when the error values have been averaged after a approximately 100 successful comparison cycles.

For practical purposes it is necessary to average the magnitude of the error over a number of samples, taking advantage of the variability of the integrator out. This provides a more consistent response. It should be noted that this response is only slightly input independent (13).

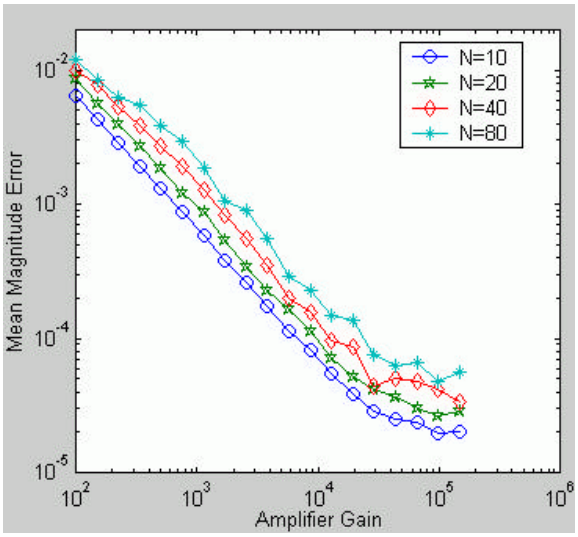

Figure 5: Errors due to finite amplifier gain.

The results show that irrespective of input values chosen, a similar result is consistently seen. For small values of gain, the error is large, tending rapidly to a small value as the gain increases, at approximately 20dB/decade.

The presence of noise, or measurement resolution plays a significant role in calculating the gain error. The measured error will always have some component due to noise and this places a lower floor on any measurement. For a given resolution, there is a maximum value of gain for which there is any discernible error. Figure 6 shows a typical example. With a measurement resolution of only 0.001 (compared to an input value of 1) it is possible to discern errors arising only due to gains less than 1000.

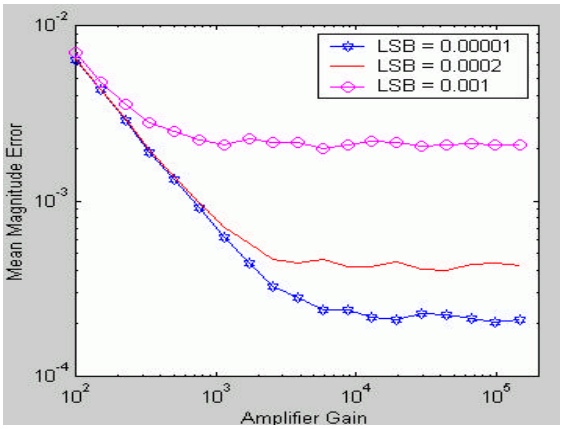

Figure 6: Finite amplifier gain errors with noise.

V CONCLUSIONS

The technique presented in this paper, though demonstrated only on the first-order modulator, is fully compatible with other sigma-delta modulator systems. It has shown itself to be capable of extracting reliable values for many of the scalar and offset errors in the system and as providing a useful indicator for poor amplifier gain. It has achieved this with a measurement period of less than 256 cumulative cycles and a limited amount of computer computation cycles. This makes it viable for use in a commercial production environment. Compared to existing techniques, this approach is highly efficient and if used prior to more extensive testing techniques could prove to be a highly cost-efficient technique for screening out bad-devices prior to more expensive tests.

Future work will involve applying this approach to higher order systems and testing this approach with real devices which may need custom circuitry to provide access to the required circuit nodes.

ACKNOWLEDGEMENTS

I would like to thank Aaron McCoy for his assistance in investigating the background issues to this problem.

REFERENCES

[1] S. Norsworthy, R. Schreier and G. Temes, “Delta-Sigma Data Converters”. *IEEE Press, ISBN 0-7803-1045-4, 1996.*

[2] J. C. Candy and G. C. Temes, Introduction of "Oversampling delta-sigma data converters", IEEE Press, ISBN 0-87942-285-8, 1992.

[3] D. Birru, and E. Roza, “Video-rate D/A converter using reduced rate sigma-delta modulation”, *Custom Integrated Circuits Conference*, Santa Clara, USA, pp. 241-244, 1998

[4] C. C. Cutler, "Transmission systems employing quantization", 1960 U.S. Patent No. 2927962, filed 1957.

[5] H. Inose and Y. Yasuda, "A unity bit coding method by negative feedback", *Proceedings of the IEEE*, Vol. 51, pp. 1524-1533, Nov 1963.

[6] R. Schreier and M. Snelgrove, "Bandpass sigma-delta modulation", *Electronic Letters*, Vol. 25, pp. 1560-1561, 1991.

[7] O. Feely, L. O. Chua, "The effect of integrator leak in Sigma-delta modulation, " *IEEE Trans. Circuits and Systems*, Vol. 38, pp. 1293-1305, 1991.

[8] R. Farrell and O. Feely, "Bounding the integrator outputs of second order sigma-delta modulators ", *IEEE Trans. Circuits and Systems II - Analog and Digital Processing*, Vol. 45, pp. 691-702, June 1998.

[9] R. Farrell, "Non-Lineary Analsys of Second Order Sigma-Delta Modulators", *Ph.D Thesis*, National Unversity of Ireland, Dublin, 1998.