ELSEVIER

Computer Standards & Interfaces 23 (2001) 103–110

COMPUTER STANDARDS & INTERFACES

www.elsevier.com/locate/csi

Analysis and simulation of a cascaded delta delta–sigma modulator

Dileepan Joseph*, Lionel Tarassenko, Steve Collins

Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, UK

Abstract

This paper analyses and simulates a delta delta–sigma modulator, the cascade combination of a delta modulator and a delta–sigma modulator. Using prediction, the delta modulator reduces the dynamic range of a signal prior to quantisation. Using noise shaping, the delta–sigma modulator cancels the error of the delta modulator. Performance is evaluated with input signals from a random noise process, composed of a high dynamic range narrow-band process and a low dynamic range wide-band process. Integrator leakage in the delta modulator and DAC mismatch between the two modulators may complicate implementation. The cascaded modulator outperforms conventional modulators of similar order when there is a degree of correlation between consecutive samples of the input process.

Keywords: Delta; Delta–sigma; Cascaded; Prediction; Noise shaping

1. Introduction

The advantages of oversampled modulation for analogue-to-digital (A/D) conversion are well known [1]. Firstly, sharp anti-aliasing is moved from the analogue to the digital domain, in which implementation is simpler. Secondly, oversampling creates a band of frequencies, above the Nyquist rate, into which quantisation noise can be shaped. Subsequent filtering and downsampling removes most of this noise. Thirdly, an increase in modulator complexity or oversampling ratio may be traded for an increase in noise shaping.

Circuit designers face the problem of maximising the performance of A/D conversion, while minimising the modulator complexity and oversampling ratio. Several architectures have been developed but first- and second-order delta–sigma modulators remain popular because of their reliability and simplicity. Higher order modulators often require multi-bit quantisers for stability, which degrades their performance and reduces their appeal. Single-bit first- and second-order modulators have been cascaded together successfully to realise higher order performance.

Fig. 1 introduces a *cascaded delta delta–sigma* (D2S) modulator, where a delta–sigma modulator converts the prediction error of a delta modulator. The D2S modulator outperforms conventional modulators, of similar complexity and oversampling ratio, for correlated input processes. If the input to an A/D converter (ADC) is correlated, then taking the difference between the input and a first-order prediction of the input reduces the dynamic range of the

* Corresponding author. Tel.: +44-1865-273805; fax: +44-1865-273905.

E-mail address: dileepan@robots.ox.ac.uk (D. Joseph).

0920-5489/01/$ - see front matter
PII: S0920-5489(01)00063-0

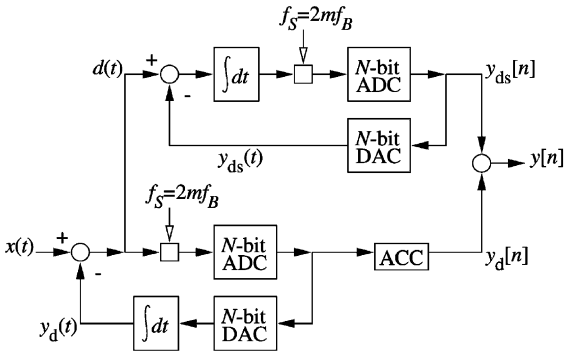


Fig. 1. A cascaded delta delta–sigma (D2S) modulator.

signal. Modulating this difference signal is less demanding than modulating the original input signal.

Section 2 describes the architecture of the D2S modulator and provides an analysis of its performance. Using simulation with a random noise process, Section 3 evaluates the modulator with respect to conventional modulators. Section 4 offers perspectives on implementation.

2. Analysis

The lower half of Fig. 1 shows a delta modulator with input $x(t)$ and output $y_d[n]$. Sampling the signal $x(t)$ at a rate f_S that is many times the Nyquist rate $2f_B$ decreases the time between consecutive samples $x[n]$ and $x[n-1]$ and increases their correlation [3]. Thus, as the oversampling ratio m increases, the dynamic range of the difference $x[n] - x[n-1]$ decreases compared to the dynamic range of $x[n]$. Delta modulation essentially quantises this difference.

The feedback signal $y_d(t)$ in the delta modulator predicts the input $x(t)$. If the prediction is less than the input, the ADC outputs a proportionate positive integer, making the D/A converter (DAC) output a positive signal to increase the prediction. If the prediction is more than the input, the ADC outputs a negative integer causing the DAC to output a negative signal to decrease the prediction. The digital accumulator (ACC) is simply an up/down counter that increments or decrements the output word using the integer updates from the ADC.

Whereas the integration in the feedback path of the delta modulator is analogue, the integration (ACC) in its output path is digital. Providing the leakage in the analogue integrator is controlled or compensated, the digital output $y_d[n]$ will approximate the analogue prediction $y_d(t)$ [4].

Fig. 1 also shows a first-order delta–sigma modulator with input $d(t)$ and output $y_{ds}[n]$. Since oversampling creates a band $f_B \le |f| \le f_S - f_B$ that is later removed by decimation, delta–sigma modulation shapes the quantisation noise into this band. The integrator in the delta–sigma modulator keeps the average value of the input $d(t)$ equal to the average value of the feedback $y_{ds}(t)$ which forces most of the quantisation noise in $y_{ds}[n]$ to higher frequencies [4].

First-order delta–sigma modulation usually involves two-level quantisation [1] although Fig. 1 shows an N-bit quantiser (the cascade combination of ADC and DAC). The analogue signal $y_{ds}(t)$ tracks the digital signal $y_{ds}[n]$ almost perfectly since D/A conversion introduces almost no error. Increasing the number of quantisation levels, however, increases the nonlinearity of the DAC and may cause harmonic distortion [4]. The choice of N is discussed further in Section 3 below.

The output $y[n]$ of the D2S modulator is the sum of $y_d[n]$ and $y_{ds}[n]$. Fig. 2 gives a linear discrete-time model [5] of the cascaded modulator, with an extra delay in the delta modulator path for synchronisation. Quantisation is modelled by the addition of white noise to the quantiser input and integration is modelled by accumulation [4].

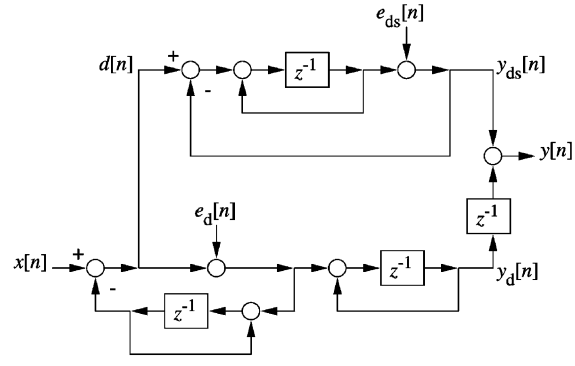


Fig. 2. Linear discrete-time model of D2S modulation.

Using the model in Fig. 2, the transfer functions of the delta modulator and the delta–sigma modulator are:

$$Y_{\mathrm{d}}(z) = z^{-1}X(z) + z^{-1}E_{\mathrm{d}}(z), \quad (1)$$

$$Y_{\mathrm{ds}}(z) = z^{-1}D(z) + (1 - z^{-1})\,E_{\mathrm{ds}}(z), \quad (2)$$

and the transfer function of the D2S modulator is:

$$Y(z) = z^{-1}X(z) + (1 - z^{-1})\,E_{\mathrm{ds}}(z). \quad (3)$$

Only the noise $E_{\mathrm{ds}}(z)$ of the delta–sigma modulator appears at the output of the D2S modulator and it is shaped towards high frequencies. The delta–sigma modulator cancels out the noise $E_{\mathrm{d}}(z)$ of the delta modulator.

Comparing Eqs. (2) and (3), it is clear that the formulation of in-band noise power σ_ε^2 for D2S modulation is identical to that of first-order delta–sigma modulation, i.e. [1]:

$$\sigma_\varepsilon^2 = \frac{\pi^2\sigma_e^2}{3m^3} = \frac{\pi^2 Q_{\mathrm{ds}}^2}{36m^3}, \quad (4)$$

where σ_{e}^2 is the power in the noise signal $e_{\mathrm{ds}}[n]$ and Q_{ds} is the step-size of the delta–sigma quantiser. Since the quantisation error $e_{\mathrm{ds}}[n]$ is assumed to be uniformly distributed from $-Q_{\mathrm{ds}}/2$ to $+Q_{\mathrm{ds}}/2$, σ_{e}^2 equals $Q_{\mathrm{ds}}^2/12$. Although D2S modulation appears to have the same in-band noise power as first-order delta–sigma modulation, there is a difference.

The step-size of a delta–sigma quantiser is chosen so that the range of reference levels can accommodate the range of the input signal plus some circulating noise [1]. Thus, denoting the maximum amplitude of the input by α, the in-band noise power for first-order delta–sigma modulation, where N is the number of quantisation bits, is:

$$\sigma_\varepsilon^2 = \frac{\pi^2\alpha^2}{9m^3(2^N-1)^2} \text{ since } Q_{\mathrm{ds}} = \frac{2\alpha}{2^N-1}, \quad (5)$$

For D2S modulation, the input to the delta–sigma modulator is $d(t)$ not $x(t)$. Thus, Q_{ds} does not depend on the dynamic range α of the input signal $x(t)$ but on the dynamic range of the prediction error $d(t)$. If the maximum slew rate $|\mathrm{d}x/\mathrm{d}t|$ of the input is denoted δ, and the sampling period is $T = 1/2mf_{\mathrm{B}}$, then the input may change by up to $\pm\delta T$ each period. Assuming the prediction is able to follow the input, the prediction error is bounded by $\pm\delta T$ and, for D2S modulation, the in-band noise is:

$$\sigma_\varepsilon^2 = \frac{\pi^2\delta^2}{36m^5f_{\mathrm{B}}^2(2^N-1)^2} \text{ since } Q_{\mathrm{ds}} = \frac{2\delta T}{2^N-1}, \quad (6)$$

For the prediction to follow the input, the range of the reference levels in the delta quantiser must cover $-\delta T$ to $+\delta T$. Therefore, the step-size of the delta quantiser is:

$$Q_{\mathrm{d}} = \frac{2\delta T}{2^N-1} = Q_{\mathrm{ds}}. \quad (7)$$

Such a choice for Q_{d} keeps the quantisation error $e_{\mathrm{d}}[n]$ within $\pm Q_{\mathrm{d}}/2$ [3]. A smaller Q_{d} may mean that the prediction will not be able to follow the input, causing a significant increase in noise when the input changes too quickly. However, using a larger Q_{d} may increase the granular noise since the prediction often alternates above and below the input during periods of slow change. Optimal choice of Q_{d} therefore depends on the probability distribution of slew rates for the input signals.

For comparison, the noise power formulae for delta modulation alone and second-order delta–sigma modulation are given in Eqs. (8) and (9), respectively [1]. If the dynamic range α of the input and the number of quantisation bits N are kept constant, the step-size Q_{ds} of the quantiser in second-order delta–sigma modulation must be larger than before (hence a smaller denominator) to accommodate a slight increase in circulating noise [1]:

$$\sigma_\varepsilon^2 = \frac{\sigma_e^2}{m} = \frac{\delta^2}{12m^3f_{\mathrm{B}}^2(2^N-1)^2} \quad \left(Q_{\mathrm{d}} = \frac{2\delta T}{2^N-1}\right), \quad (8)$$

$$\sigma_\varepsilon^2 = \frac{\pi^4\sigma_e^2}{5m^5} = \frac{\pi^4\alpha^2}{15m^5(2^N-2)^2} \quad \left(Q_{\mathrm{ds}} = \frac{2\alpha}{2^N-2}\right). \quad (9)$$

Doubling the oversampling ratio m in either first-order delta–sigma or delta modulation decreases

the noise power by a factor of eight and, hence, increases the signal-to-noise ratio (SNR) by 9 dB. Doubling m in either second-order delta–sigma or D2S modulation decreases the noise power by a factor of 32 and, hence, increases the SNR by 15 dB. Whereas the dynamic range α of the input does not limit the performance of delta modulation and D2S modulation, the maximum slew rate δ does. The reverse is true for delta–sigma modulation.

3. Simulation

Delta, first- and second-order delta–sigma and D2S modulation were simulated in Matlab. Test signals $x(t)$ were realisations of a process $X(t)$, shown in Fig. 3, composed of two *white* processes $A(t)$ and $B(t)$, both generated from uniformly distributed random numbers. A white process is uncorrelated and has a uniform power spectral density (PSD) [3]. Process $A(t)$ has a bandwidth $|f| \le f_A$ with a PSD equal to a and process $B(t)$ has a bandwidth $|f| \le f_B$ with a PSD equal to b. Because they are statistically independent, the PSD of the sum of the two processes equals the sum of their PSDs. The Nyquist rate of $X(t)$ is $2f_B$, since $f_B \ge f_A$, but $X(t)$ is oversampled at $f_S = 2mf_B$.

Three simulations were undertaken, varying the oversampling ratio m, the bandwidth ratio f_B/f_A and the power ratio af_A/bf_B (the power ratio corresponds to the ratio of dynamic ranges). For each configuration, 10 sample signals $x(t)$ were generated and the SNR was averaged over these realisations. The D2S modulator was implemented with two one-bit (N equals one) quantisers (with Q_d equal to Q_{ds}).

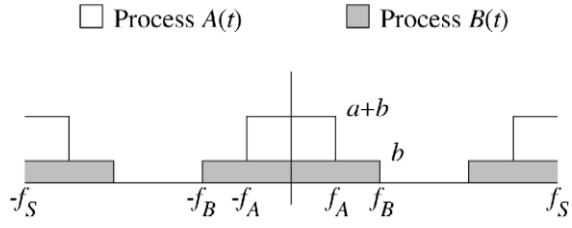


Fig. 3. Power spectral density of a test process $X(t)$, constructed for simulation purposes.

To keep the total bit-rate constant, the delta modulator and delta–sigma modulators (first and second-order) were implemented with two-bit (N equals two) quantisers.

Fig. 4 shows the SNR performance of the four modulators vs. the oversampling ratio. For this simulation, both the bandwidth and the power ratio are set to a value of 64, meaning that process $A(t)$ has eight times the dynamic range within $1/64$ times the bandwidth of process $B(t)$. Process $X(t)$ is then a slowly varying signal of high dynamic range superimposed on a quickly varying signal of low dynamic range. There is, therefore, sufficient correlation between consecutive samples for the predictive modulators (delta and D2S) to outperform the other two.

As expected, Fig. 4 shows that SNR improves with increasing m. The SNR of the first-order modulators (delta–sigma and delta) increases by 9 dB/octave, while the second-order modulators (delta–sigma and D2S) improve by 15 dB/octave, consistent with the theory. Except at low oversampling ratios, D2S modulation gives the best results. First-order modulation is better than second, for low m values, because the constants in Eqs. (5) and (8) are smaller than the ones in Eqs. (9) and (6), respectively.

For an oversampling ratio of 32 and a power ratio of 64, Fig. 5 plots the SNR of the modulators vs. the bandwidth ratio. The bandwidth ratio f_B/f_A corresponds to the separation, in rate of variation, between the processes $A(t)$ and $B(t)$. The leftmost point, in Fig. 5, corresponds to a unit ratio, where f_A equals f_B. At this point, $X(t)$ is the sum of two white processes of equal bandwidth, but different average power. Such a process is relatively uncorrelated so the delta–sigma modulators are slightly better than the predictive modulators. However, the whiteness of $X(t)$ decreases quickly as the bandwidth ratio increases, which increases correlation and allows each predictive modulator to surpass the delta–sigma modulator of similar order.

The performance of the delta–sigma modulators is not affected much by the bandwidth ratio. Increasing the ratio does not change the maximum amplitude of $X(t)$ since the dynamic range of $A(t)$ and $B(t)$ are unaffected. However, the maximum slew rate of $X(t)$ decreases as f_B/f_A increases since the rate of variation of $A(t)$, which accounts for most of the power in $X(t)$, decreases. Thus, the SNR of

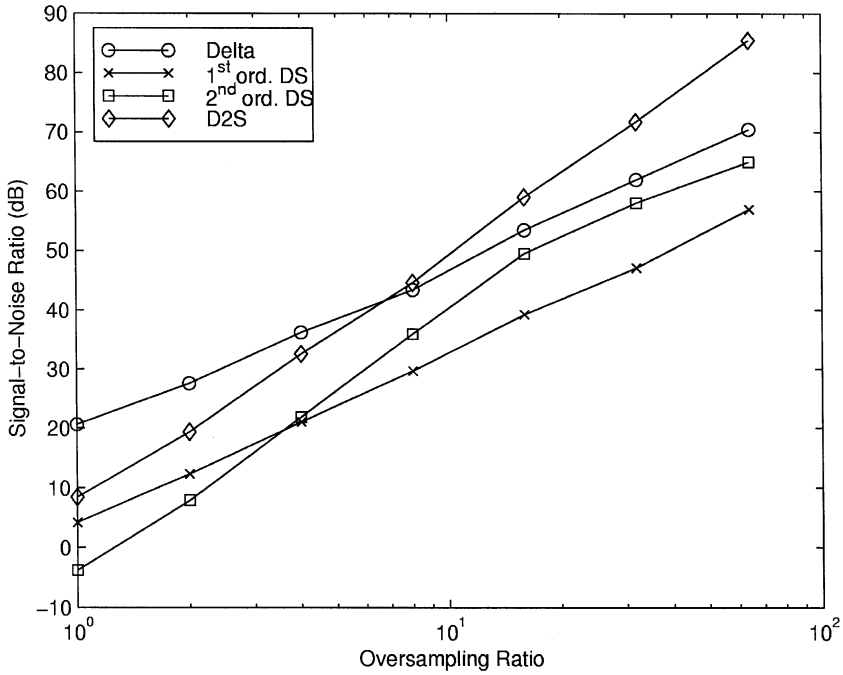


Fig. 4. Modulator signal-to-noise ratio vs. the oversampling ratio of the test process.

predictive modulation increases with increasing bandwidth ratio. However, the SNR eventually saturates since $A(t)$ behaves like a large DC signal with a random amplitude for high bandwidth ratios (i.e.

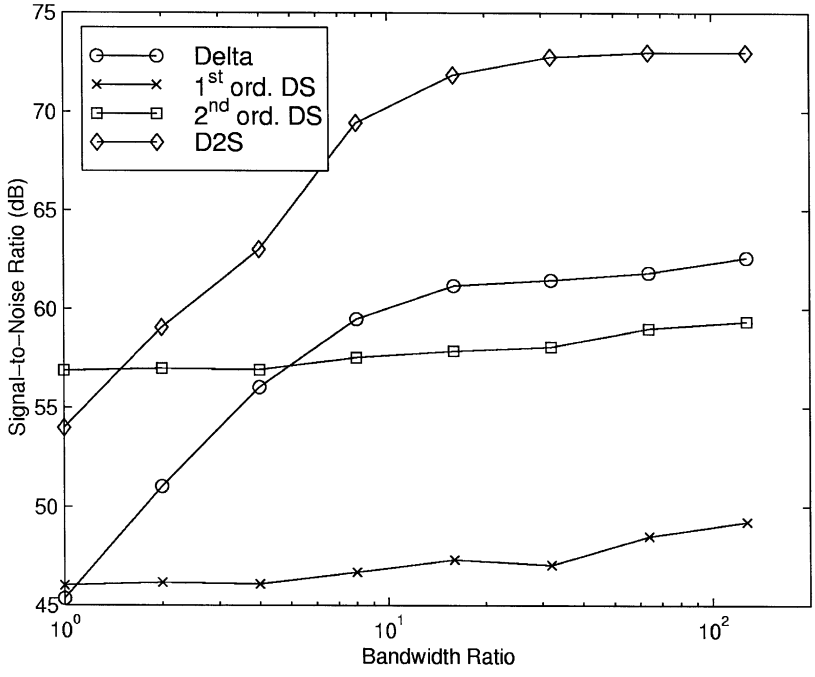


Fig. 5. Modulator signal-to-noise ratio vs. the bandwidth ratio of the test process.

above 16), which limits the maximum slew rate of $X(t)$ to the maximum slew rate of $B(t)$.

As shown in Fig. 4, second-order surpasses first-order modulation once m is sufficiently high. Each second-order modulator (delta–sigma and D2S) is better than its first-order counterpart (delta–sigma and delta) because of an increase in or, simply, the use of noise shaping. This result is also evident in Fig. 6, which charts the SNR performance of the four modulators against the power ratio af_A/bf_B of the test process. The oversampling ratio was 32 for this simulation and the bandwidth ratio was 64.

A unit power ratio, at the centre of Fig. 6, corresponds to $A(t)$ and $B(t)$ having equal dynamic range. D2S modulation is then slightly worse than second-order delta–sigma modulation; however, delta modulation is better than first-order delta–sigma modulation. At the left end of the graph, $A(t)$ has one-tenth the dynamic range of $B(t)$. D2S modulation is then worse than second-order delta–sigma modulation, while delta modulation and first-order delta–sigma modulation are similar. However, when $A(t)$ has a larger dynamic range than $B(t)$, each predictive modulator is much better than the delta–sigma modulator of similar order.

The performance of the delta–sigma modulators does not really change with an increasing power ratio since a change in the dynamic range of $X(t)$ affects both the signal power and noise power equally, leaving the SNR results of the delta–sigma modulators unchanged. However, the correlation in $X(t)$ increases with an increasing power ratio af_A/bf_B since more power is allocated to the slowly varying component $A(t)$. As a result, the SNR results of predictive modulation improve with increasing power ratio.

4. Implementation perspectives

There are principally two difficulties that may arise when implementing this architecture: integrator leakage and DAC mismatch. Leakage of the integrator in the feedback path of the delta modulator will cause errors as the output $y_d[n]$ of the digital integrator will not match the output $y_d(t)$ of the analogue integrator. However, leakage of the integrator in the feedforward path of the delta–sigma modulator is not significant, as the feedback action will correct

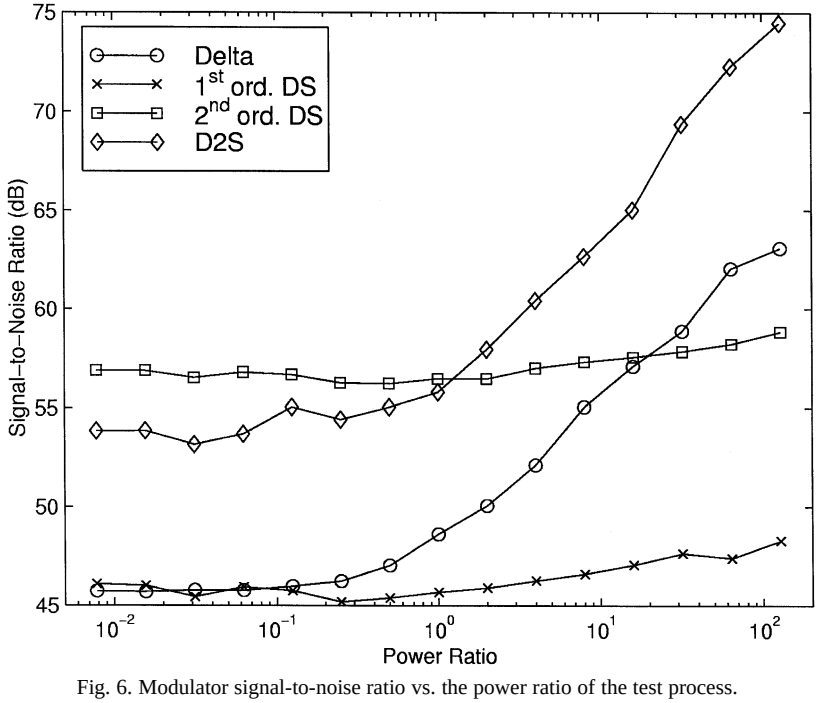


Fig. 6. Modulator signal-to-noise ratio vs. the power ratio of the test process.

the error. Leakage may be reduced by careful design of the integrator.

A small constant current discharging the integrator may cause leakage [2]. This leakage will add a constant ramp to the output signal that may be calibrated out on the digital side. Leakage may also be resistive, where a percentage of the charge stored on the integration capacitor is lost each sampling period [2]. The error, for this leakage, will depend on the absolute voltage across the integrator and is more difficult to calibrate out. The simplest way to solve leakage problems is to periodically reset the delta modulator so that the analogue and digital integrators are in equivalent states [2]. Given that leakage is typically slow, resetting the modulator would have little effect on overall performance.

A second difficulty arises if the output voltage levels of the DACs, found in the delta and delta–sigma feedback loops, do not match, as required by Eq. (7). Mismatched DACs can be modelled by a gain in the feedback path of the delta modulator. A mismatch will prevent the delta–sigma modulator from cancelling the error $E_d(z)$ of the delta modulator fully. This problem is not unique to the D2S modulator, but is a problem of all cascaded modulators, including the popular combination of two delta–sigma modulators. The degree of matching necessary to achieve a particular performance can be readily modelled and a circuit must be built to match within the required tolerances [1].

5. Conclusion

This paper has investigated circuitry that improves A/D conversion of oversampled signals for a certain type of input. Although the cascaded delta delta–sigma (D2S) modulator discussed here is similar to the oversampled modulator mentioned in Gaboury and Harjani [2], the analysis and simulation discussed here gives a more complete explanation of the architecture and of the reasons for its performance. If the main benefit of oversampled modulation is that it permits a trade-off between several parameters of A/D conversion, then this paper shows how the D2S modulator trades generality for performance.

Delta modulation uses prediction to reduce the dynamic range of a signal before quantisation. Delta–sigma modulation shapes the quantisation noise spectrum so that less noise falls in the Nyquist band. This paper has shown that prediction and noise shaping may be usefully combined in a D2S modulator, in which a delta–sigma modulator cancels out the error of a delta modulator. Simulation results agree with theoretical analysis in showing that D2S modulation outperforms first- and second-order delta–sigma modulation, and delta modulation, when converting signals composed of a high dynamic range narrow-band process and a low dynamic range wide-band process.

Acknowledgements

DJ gratefully acknowledges the support of the Natural Sciences and Engineering Research Council of Canada.

References

[1] J.C. Candy, G.C. Temes, Oversampling methods for A/D and D/A conversion. Oversampled Delta–Sigma Data Converters: Theory, Design and Simulation. IEEE, New York, 1992, pp. 1–29.

[2] M. Gaboury, R. Harjani, Delta–sigma–delta modulator: an autoranging A/D converter. IEEE International Symposium on Circuits and Systems. IEEE, New York, 1997, pp. 401–404.

[3] S. Haykin, Communication Systems. 3rd edn., Wiley, New York, 1994.

[4] W.L. Lee, C.G. Sodini, A Novel Higher-Order Interpolating Modulator Topology for High Resolution Oversampling A/D Converters, MSc thesis, Massachusetts Institute of Technology, 1987.

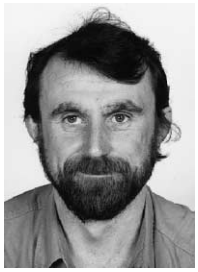
[5] A. Papoulis, Circuits and Systems—A Modern Approach. Holt, Rinehart and Winston, Chicago, 1980.

Dileepan Joseph received a BSc in Computer Engineering from the University of Manitoba, Canada, in 1997. He graduated with distinction, winning the University Gold Medal in Engineering and the Governor General's Silver Medal. Presently, he is pursuing a DPhil in Engineering Science at the University of Oxford under a fellowship from the Natural Sciences and Engineering Research Council of Canada. His doctoral research concerns logarithmic CMOS image sensors but he worked previously on distributed genetic algorithms, phase-locked-loop oscillators and oversampled modulation.

Professor Lionel Tarassenko, MA, DPhil, FREng, FIEE was born in Paris in 1957. He gained the degrees of BA in Engineering Science in 1978, and DPhil in Medical Engineering in 1985, both from Oxford University. After graduating, Lionel Tarassenko worked for Racal Research on the development of digital signal processing techniques, principally for speech coding. He returned to Oxford University at the end of 1981 to study for a DPhil in medical electronics. He then held a number of positions in academia and industry, before taking up a University Lecturership in Oxford in 1988. Since then, he has devoted most of his research effort to the development of neural network techniques and their application to signal processing, diagnostic systems, and parallel architectures. Over the last 10 years, he has received 3.5 million pounds in research grants from EPSRC, DTI, the Wellcome Trust, the EU and directly from industry. In addition, he is a Principal Investigator in the recently awarded IRC "From Medical Images and Signals to Clinical Information" (9 million pounds from EPSRC and MRC).

Professor Tarassenko has led three projects all the way from concept to commercial exploitation: the Sharp LogiCook Microwave Oven (the first of Sharp's products to be developed outside Japan and the first microwave oven to incorporate neural network control), the QUESTAR sleep analysis system and QUINCE, a jet engine diagnostic system for Rolls-Royce. QUESTAR was awarded a British Computer Society Medal in 1996.

Professor Tarassenko is the author of 70 refereed publications and 60 conference papers, mostly in the field of neural networks, including two books. He has been the holder of the Chair in Electrical Engineering at Oxford University since October 1997. He was elected to a Fellowship of the Institution of Electrical Engineers (IEE) in 1996, when he was also awarded the IEE Mather Premium for his work on neural networks, and to a Fellowship of the Royal Academy of Engineering in 2000. He is a non-executive director of ThirdPhase and a founder director of Oxford BioSignals.



Steve Collins received a BSc in Theoretical Physics from the University of York in 1982 and a PhD from the University of Warwick in 1986. From 1985 until 1997 he worked at the Defence Research Agency on various topics including the origins of $1/f$ noise in MOSFETs and analogue information processing. From 1997 he has been at the University of Oxford were he has continued his interest in smart imaging sensors and non-volatile analogue memories.