Analysis of Feedback Predictive Encoder Based ADCs

ABSTRACT:

Purpose- The purpose of this paper is to present analysis of the feedback predictive encoder based ADC (Analog-to-Digital Converter).

Design/methodology/approach- The use of feedback predictive encoder based ADCs presents an alternative to the traditional two stage pipeline ADC by replacing the input estimate producing first stage of the pipeline, with a predictive loop that also produces an estimate of the input signal.

Findings- The overload condition for feedback predictive encoder ADCs is dependent on input signal amplitude and frequency, system gain and filter order. The limitation on the practical useable filter order is set by limit cycle oscillation. A boundary condition is defined for determination of maximum useable filter order. In a practical implementation of the predictive encoder ADC, the time allocated to the key functions of the gain stage and loop quantizer leads to optimisation of the power consumption.

Originality/value- This paper presents a methodology to optimise the bandwidth of predictive encoder ADCs. The overload and stability conditions may be used to determine the maximum input signal bandwidth for a given loop quantizer. Optimisation of power consumption based on the allocation of time between the gain stage and the SAR ADC operation is investigated. The lower bound of power consumption for this architecture is estimated.

Paper Type- Research Paper
1. Introduction & motivation.

Two step ADCs (Analog-to-Digital Converters) quantize a signal in a series of sequential steps, with the 1st stage quantizer resolving a set number of bits $N_1$ and passing a gained residue signal to the 2nd stage which is only required to consist of a quantizer assumed to resolve $N_2$ bits. The block level model of the two stage pipelined ADC is detailed (Figure 1a). Typically a sample and hold function precedes the 1st stage quantization and gain functions (Hernes et al., 2007). This input sample and hold provides time synchronization between the high quality analog sample and the sample processed by the 1st stage quantizer. The exact architecture of the 2nd stage is unimportant in the analysis of the two step conversion process but does affect the latency of the conversion. With a flash or high speed SAR (successive Approximation Register) ADC providing fast conversion (Lee and Flynn, 2010, Van der Goes et al., 2014), lower latency for the total conversion may be achieved. The additional number of bits is determined by the gain $G$ between the stages, therefore the total number of bits resolved by the two step ADC $N_{tot}$ is given by,

$$N_{tot} = \log_2(G) + N_2$$

where any additional bits resolved in the first stage provide redundancy, the following inequality must hold

$$\log_2(G) \leq N_1.$$  

The purpose of the path formed by the 1st stage quantizer and DAC (Digital-to-Analogue Converter) of the 1st stage is to provide an estimate of the input signal. However in the conversion process there is no requirement as to how the estimate is to be generated. An alternative predictive encoder based approach to generating the input signal estimate is detailed in Harris (1987). In this approach (Figure 1b), the predictive filter may be connected in feedback around the 2nd stage quantizer to produce the estimate of the input signal. Firstly the predictive estimation allows removal of the 1st stage flash path, however a DAC function is still required to convert the prediction estimate to an analog signal. As in the two step ADC approach, the DAC providing the input signal estimate must be accurate to the overall system level accuracy. Secondly the predictive encoder allows elimination of the front end sample and hold as the synchronization between analogue input signal and the estimate is replaced by the predictive estimate. This predictive signal compensates for the sample delay in the processing through the loop quantizer. However the delay compensation provided by the predictor is bandwidth limited which limits the overall input bandwidth of the predictive encoder based ADC approach. This approach is therefore classified as oversampled with higher order predictors increasing the theoretical input signal bandwidth.

The aim of this paper is the presentation of an analysis for determination of the optimal selection of prediction filter order for maximum input bandwidth. The presented approach determines the input bandwidth limitations from the quantizer no-overload condition. This approach is similar to methods used to determine input-amplitude stability from the class of multi-bit converters (Schreier and Temes, 2005), but requires modification for the predictive encoder case. This paper is organized as follows; Section 2 compares the relationship of the predictive encoder to the sigma delta modulator. Section 3 details the design of digital prediction filters for use in the predictive encoder ADC. Section 4 presents the analysis of feedback predictive encoder based ADC and determines the no-overload condition for maximum input frequency. Consideration is given to the limit cycle oscillatory behavior of the feedback configuration in Section 5. System level simulations are detailed in Section 6. A practical implementation for the predictive encoder based ADC is presented in Section 7. The lower bound for power consumption for this implementation is discussed and a trade off in time allocation to circuit blocks allows power consumption to be optimized. Finally conclusions are presented in Section 8.

2. Relationship to Sigma Delta converters.

The predictive encoder based ADC is a member of the class of linear predictive and noise shaped coders. An encoder (e.g. predictive encoder) is typically used to convert a high resolution sampled signal into a coarsely quantized high frequency signal that is robust to
noise or interference during transmission. The encoders (Zrilic, 2006) and their associated ADCs are detailed in Figure 2. An encoder also has an associated ADC which converts the input signal into a finely quantized output signal which may be used for digital signal processing. The relationship between the predictive encoder and the sigma delta modulator illustrated by the example shown in Figure 3, for a 2\textsuperscript{nd} order system based on a cascade of integrators. It is noted that the difference in the block diagram between the sigma delta modulator, the predictive encoder and the hybrid system is the position of the input signal \( S \). Note that in both the sigma delta modulator and predictive encoder the quantization noise is high pass filtered. The result of the positioning of the input signal is that in the predictive encoder the input signal is also high pass filtered, whereas in the sigma delta modulator the input signal is delayed. Note that a hybrid predictive encoder/sigma delta modulator structure can be produced as shown in Figure 3c. Due to the similarity of their structures the input amplitude conditions of predictive encoders and sigma delta modulator may be similarly determined in terms of the no-overload condition at the quantizer (Schreier and Temes, 2005, Lokken et al., 2006). The determination of the no-overload condition for the predictive encoder ADC is detailed in Section 4.

The sigma delta modulator ADC is well known as the most popular type of oversampled data converter with excellent figure of merit and power consumption reaching thermal noise limits (Dong et al. 2014). It is important to note however that when the predictive encoder and sigma delta modulator based are utilized in an ADC configuration (Figure 2c,2d), there is a significant difference in the manner in which additional resolution above that available in the loop quantizer is generated. The sigma delta modulator ADC takes advantage of the high pass filtering of the quantization noise to increase the effective number of bits in the in-band frequency region, a low pass filter is required after the modulator to remove the high pass filtered quantization noise. In the predictive encoder ADC case analogously to the two step ADC (Figure 1a,b), a gain of the residue \( U \) between input signal \( S \) and input estimate \( V \) is used to generate additional bits of resolution. Increasing noise shaping filter order in the sigma delta modulator increases converter resolution, this may be also achieved by cascading lower order stages (Ali Sadat Noori et al. 2016). While in the predictive encoder based ADC increased predictive filter order increases system bandwidth.

While prediction filters may be implemented as a cascade of integrators in the analog domain, modern CMOS (Complementary Metal Oxide Semiconductor) technologies make the design of high order prediction filters feasible in the digital domain.

### 3. Prediction filters.

In order to compensate for the sample time delay which is equivalent to a linearly increasing phase lag between the input and estimated input, a prediction filter is connected in feedback around the 2\textsuperscript{nd} stage quantizer now referred to as the loop quantizer. The ideal prediction filter will have a flat magnitude and increasing phase over the Nyquist bandwidth. The digital prediction filters may be derived as an extrapolated form of the Newton’s series or more generally from arbitrary high pass filters such as the Chebyshev form (Vaidyanathan,2008). In general the digital prediction filter \( H_p(z) \) has an associated high pass filter \( H_{hp}(z) \) where

\[
H_p (z) = 1 - H_{hp} (z).
\]  

(3)

Realizable prediction filters have a limited theoretical bandwidth over which there is an approximately flat magnitude and linearly increasing phase. The prediction filter will approach its theoretical maximum bandwidth with the selection of higher filter orders. The basic Newton’s series predictor (4) and its modified form (5) for prediction of a sample \( x(t) \) at time \( t \), may be expressed in terms of binomial expansions for a given filter order \( n \) and \( \alpha \) is a constant in (5). These forms have finite bandwidths of \( f_s/6 \) and \( f_s/4 \) respectively,

\[
s(t) = \lim_{n \to \infty} \sum_{k=1}^{n} (-1)^{k+1} \binom{n}{k} s(t - kT),
\]

\[
BW_{EFF} \leq \frac{f_s}{6}.
\]  

(4)
\[ s(t) = \lim_{n \to \infty} \sum_{k=1}^{n} (-1)^{k+1} (\cos \pi n)^k \binom{n}{k} s(t - kT), \]

\[ \text{BW}_{\text{EFF}} \leq \frac{L_s}{4}. \]  

(5)

Figure 4a details the frequency response of the predictive filter of form (5) with order \( n = 11 \). The finite region of flat magnitude corresponds to the finite stop band of the associated high pass filter shown in Figure 4b. The predictive filter and its associated high pass filter are related by (3). The non-flat filter magnitude response limits the effective input bandwidth of the ADC. The parameter \( a < 0.5 \) determines the theoretical maximum bandwidth of the prediction filter of the form (5). Increasing the prediction filter order would allow the expected theoretical input bandwidth of the prediction filter to be approached.

4. Predictive encoder ADC quantizer no-overload analysis.

The predictive encoder ADC resolution is determined by the choice of the inter-stage gain \( G \) and the number of bits in the loop quantizer \( N_z \). Referring to the discrete time model in Figure 5 (ignoring saturation) the Z domain input to the loop quantizer is given by

\[ U(z) = G \cdot H_{hp}(z)S(z) - H_p(z)Q_z(z). \]  

(6)

The signal at the input to the quantizer contains a high pass version of the input and predictive filtering of the quantization noise. Therefore if the input signal is within the stop band of the high pass filter the \( H_{hp}(z)S(z) \) term will be small, if however the input signal enters the pass band of the high pass filter the residue signal amplitude will be increased. The \( H_{hp}(z)S(z) \) will be referred to as the prediction error. As this signal is passed through the gain block to the quantizer, input frequency as well as input amplitude dependent overload of the loop quantizer can occur. The contribution to signal amplitude at the loop quantizer from filtered quantization noise \( H_p(z)Q_z(z) \) must also be considered. If a very coarse quantization is performed, a large amount of quantization noise which is high pass filtered is produced, substantially reducing the signal range available for the input signal.

In order to quantify the maximum input amplitude and frequency for single tone sinusoids, the no-overload condition at the loop quantizer must be determined. The variance of the quantization noise power at the loop quantizer is determined as

\[ \sigma_q^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_q(\omega)|H_p(j\omega)|^2 d\omega, \]  

(7)

where \( S_q(\omega) \) is the power spectral density of the quantization noise, and is equal to

\[ S_q(\omega) = q^2 / 12. \]  

(8)

From Parseval’s Theorem (7) can be written as

\[ \sigma_q^2 = \frac{q^2}{12} \sum_{n=0}^{\infty} |h_n|^2, \]  

(9)

where \( h_n \) are the coefficients of the FIR (Finite Impulse Response) prediction filter \( H_p \).

The variance at the loop quantizer due to the sinusoidal input of amplitude \( A \), at a single frequency \( \omega_c \) is determined as

\[ \sigma_s^2 = G^2 \frac{A^2}{2} |H_{hp}(j\omega_c)|^2. \]  

(10)

This term \( \sigma_s^2 \) may also be referred to as the prediction error power. Assuming an allowable maximum total signal power at the quantizer input \( (A = V_{\text{ref}}/2) \) a condition for maximum frequency \( \omega_c \) of the input sinusoid with amplitude \( A \) for a given choice of prediction filter, will occur when the equality given in (11, 12) is satisfied.
\[(V_{\text{ref}}/2)^2/2 \geq \sigma_s^2 + \sigma_q^2, \quad (11)\]
\[(V_{\text{ref}}/2)^2/2 \geq G^2 \frac{\sigma_s^2}{2} |H_{dp}(j\omega_c)|^2 + \frac{1}{12} \sum_{n=0}^{\infty} |h_n|^2. \quad (12)\]

In practice with \(N_2 \geq 5\) the noise term is small and the bandwidth is dominated by the prediction error term.

Figure 6 details an example of the calculated prediction error power \(\sigma_s^2\) and quantization noise power \(\sigma_q^2\) for a predictive encoder based ADC. The predictive error power is observed to increase with frequency to become the dominant error source. The associated SINAD (Signal to Noise and Distortion Ratio) produced by the predictive encoder based ADC system implemented in MATLAB simulink is also shown. It is noted that as the normalised predictive error power approaches 0dB the SINAD begins to reduce. The predictive error power is normalised relative to a full scale signal at the input of the quantizer.

The inequality of equation (12) is used to determine the theoretical bandwidth performance of the predictive encoder based ADC. The effect of increasing filter order on bandwidth (for various input amplitudes) is shown in Figure 7, it is noted that increasing filter order increases the input signal bandwidth. It is also observed that the bandwidth may be further extended by reducing input signal amplitude. It is noted that the maximum achievable bandwidth is set by the predictive filter type as detailed in Section 3.

As the gain affects the magnitude of the predictive error at the quantizer input, the trade off between bandwidth and resolution (due to increased gain) for a fixed loop quantizer resolution is detailed in Figure 8. It is noted that increasing the gain and hence resolution reduces the bandwidth of the system. This theoretical analysis based on the no-overload condition for predictive encoders suggests that the highest possible order prediction filter should be chosen as it maximises system bandwidth, however the use of high order filters increases the likelihood of limit cycle behaviour occurring and this is discussed in the following section.

5. Stability and non-linear behavior.

In a physically realizable circuit implementation of the predictive encoder ADCs circuit as shown in Figure 5, a delay is required in the loop. This may be accomplished by ensuring the integrator \(L_n(z)\) formed by the prediction filter \(H_{pn}(z)\) (order \(n\)) and the positive feedback loop is a delaying integrator. The transfer function of the integrator loop is given by,

\[L_n(z) = H_{pn}(z)/(z - H_{pn}(z)).\] \((13)\)

In the linear region of operation of the predictive encoder ADCs (i.e. no saturation) the open loop forward gain and phase is determined only from the integrator frequency response and quantizer delay. There is no contribution from the ideal gain blocks as the \(G\) and \(1/G\) stages cancel in the no-saturation case. Figure 9 details the gain and phase plots for the predictive encoder ADC system operating in open loop, for various orders of prediction filter of the form \((5)\). It is observed that for high order filters the phase lag exceeds -180\(^\circ\), potentially enabling limit-cycle oscillation to occur during normal operation. A stable limit cycle oscillation may occur when the conditions for limit cycle oscillation; -180\(^\circ\) phase shift and magnitude = 0dB are met. Since the integrator gain \(|L(j\omega)|\) is typically in excess of 0dB (Figure 9), additional gain attenuation must occur in the loop to maintain oscillation. The presence of the saturation block within the loop (Figure 5) can provide additional compression of the signal at the oscillation frequency. In the case of sinusoidal type limit cycle oscillation describing function method may be used for the non-linear saturation occurring in the loop (Condon and Hayes 2014). The describing function (Franklin et al., 2015) of a gain stage with saturation for an input signal of amplitude \(A\) is given by \((14)\).
\[
N(A) = \left\{ \begin{array}{ll}
\frac{2\alpha}{\pi} \sin^{-1} \left( \frac{\pi}{G \cdot A} \right) + \left( \frac{\pi}{G \cdot A} \right) \sqrt{1 - \left( \frac{\pi}{G \cdot A} \right)^2} & |G \cdot A| > T \\
G & |G \cdot A| \leq T
\end{array} \right.
\]  

(14)

The predictive encoder ADC open loop forward path gain \(|H_f(j\omega)|\) in the case saturation at the output of gain block is given by

\[
|H_f(j\omega)| = N(A) \cdot |L_n(j\omega)| / G.
\]

(15)

Equation (15) may be used to determine the amplitude of the limit cycle oscillation by setting \(|H_f(j\omega)|\) to unity. The frequency of oscillation is that at which there is a \(-180^\circ\) phase shift \(\omega = \omega_{-180^\circ}\) and the required gain/saturation block input amplitude \(A = A_1\) to maintain oscillation

\[
|H_f(j\omega_{-180^\circ})| = 1 = N(A_1) \cdot |L_n(j\omega_{-180^\circ})| / G.
\]

(16)

The value of \(A_1\) is not easily determined analytically from (16) as it is a function of the describing function \(N(A)\) given by (14). A graphical method is used to determine the oscillation amplitude. Figure 10 plots the describing function \(N(A)\) for a range of input amplitudes. These are multiplied by the value of the integrator loop gain at the expected oscillation frequency \(|L_n(j\omega_{-180^\circ})|\) for various values of the prediction filter order. The particular oscillation amplitude corresponds to where the curve for each filter order crosses the unity gain value. Figure 11 plots the theoretical limit cycle oscillation amplitudes against the measured values from a MATLAB Simulink model of Figure 5. It is noted that for the filter orders \((n = 11, 12, 13)\) that the measured oscillation frequency is the same (0.25π). This effect is observed as the exact oscillation frequency also depends on quantizer uncertainty (Engelen and Sarroukh, 1997). The saturation block effectively behaves as a 2 level quantizer as the input amplitude to the gain/saturation block increases. For frequencies that are rational multiples of the sample rate quantizer uncertainty can cause an additional phase shift. This additional phase shift adds to the phase contributions from the integrator and delay and allows oscillation to occur at frequencies that are rational multiples of the sample rate. These measured oscillation frequencies are close to the expected oscillation frequencies as the phase shift introduced due to quantizer uncertainty is small.

For a given prediction filter order \(n\), a boundary condition for no limit cycle oscillation may exist which is determined from (15). For no oscillation to occur at the expected oscillation frequency, the gain/saturation block should not produce sufficient gain or attenuation such that \(|H_f(j\omega_{-180^\circ})| = 1\) for all possible input amplitudes \(A\) (to the gain saturation block). The no limit cycle oscillation boundary condition will be satisfied when

\[
1 > N(A) \cdot |L_n(j\omega_{-180^\circ})| \quad \forall \ A \in [0, \infty) .
\]

(17)

From Figure 10, stable predictive encoder based ADCs are identified by a magnitude \(|H_f(j\omega)|\) not exceeding the unity gain line for all gain saturation block input amplitudes.

In lower order predictive encoder ADCs structures (e.g. \(n = 7, \alpha = 0.4\)) the natural oscillation will not occur due to sufficiently large phase margin in the integrator frequency response. However during the start-up phase or during transient inputs a large signal may occur at the input subtraction node that is sufficient to saturate the gain stage output. This mode of operation may occur in higher order filters with large signal inputs also. In the predictive encoder ADCs case saturation is followed immediately by a sampling stage (loop quantizer). The result of this sampling is to alias harmonics of the signal that are greater than the Nyquist frequency to in-band frequencies, which are then gained by the integrator. The predictive encoder ADCs structure therefore may enter a mode of tracking to the aliased harmonic of
the input. An example of this mode of operation is detailed in Figure 12. This mode of operation is highly non-linear and the SINAD produced by the predictive encoder ADCs is extremely poor. This behavior can easily be forced to occur at large input amplitudes or at input frequencies above the expected bandwidth of the predictor. It is important therefore to limit input signal amplitude and to provide sufficient anti-alias filtering to the predictive encoder ADCs structure to prevent high frequency inputs triggering this mode of operation.

6. Block level time domain simulations.

A time domain MATLAB Simulink model of the predictive encoder was used to compare theoretical results of Section 4 and 5 against simulation. Figure 13 details performance of the ADC with increasing filter order for an input amplitude of $0.9V_{ref}/2$ compared against the predicted theoretical result. The theoretical maximum input frequency based on the no-overload condition of (12) suggests that very high order predictive filters are a preferential choice for the predictive encoder based ADC. In comparison the simulated performance of the predictive encoder shows a roll off in maximum input bandwidth at a normalized frequency of $0.19f_s/2$. This result is explained by the fact that higher order filters may be conditionally stable and prone to limit cycle oscillation as detailed in Section 5.

In order to design for maximum bandwidth at a given resolution the results determined in Section 4 and Section 5 are applied. As previously seen in Figure 8 increased system gain reduces the achievable bandwidth. Therefore when designing a system for a fixed resolution, the highest possible number of bits should be chosen for the loop quantizer in order to reduce the gain required in the system. Secondly the highest possible order prediction filter that yields a stable system should be chosen. This upper limit is determined as detailed in Section 5, where the boundary condition (17) defines the maximum useable filter order. Figure 14 details a comparison between two possible 12bit systems, where a 9bit quantizer is available with a gain of 8 and a 6bit quantizer is available with a gain of 64. A high prediction filter order of 8 is chosen for both cases, which has been determined from the analysis in Section 5 to be the maximum achievable prediction filter order (form (5)) that yields a stable loop (Figure 10). It is noted from Figure 14 that the system with the larger loop quantizer ($N_\ell = 9$) has a significant bandwidth advantage. In a circuit level implementation consideration must be given to the difficulty of designing low latency quantizers with large numbers of bits and also to high speed gain blocks. Considerations for the circuit level implementation of a predictive encoder are now discussed.

7. Circuit implementation considerations.

From the simulation results of the previous section it has been determined that a predictive encoder ADC system with low values of gain and a large number of bits resolved by the quantizer has a significant bandwidth advantage. Figure 15 presents a discrete time realization of such a predictive encoder based ADC using CMOS switched capacitor technology. The estimated power consumption, implementation issues and advantages are discussed for each sub-block. An example system of 12bit resolution operating at 100MHz is assumed in order to estimate power consumption and allow comparison to state of the art conventional pipeline and SAR ADCs.

(a) Switched capacitor gain stage

The input voltage is initially sampled and gained by a switched capacitor gain stage. Subtraction of the predicted input voltage may also be achieved in this stage by utilizing a multiplying DAC type stage to combine the sampling capacitor and DAC capacitors as in a standard switched capacitor pipeline ADC (Mehr and Singer, 2000). The input gain stage must be linear to the required resolution (e.g. 12bits) of the ADC and both the sampled thermal noise and the amplifier thermal noise should be designed for levels below Signal to Noise Ratio ($SNR_Q$) for the quantization noise. For an $N_{\text{tot}}$ bit ADC the Root Mean Square (RMS) signal voltage is $2^{N_{\text{tot}}}$ times the RMS quantization voltage. Steensgaard (2004) gives
the Signal to Noise Ratio with respect to thermal noise sources ($SNR_T$), for a fully differential input gain stage as

$$SNR_T = \frac{A}{\sqrt{\frac{kT}{4C_{\text{sample}}} + \frac{8}{3g_{m1}T_{BW}}}}.$$  \hspace{1cm} (18)

Where $g_{m1}$ is the transconductance of the amplifier first stage, $T_{BW}$ is the gain stage time constant, $kT = G/2\pi \cdot GBW$, $k$ is Boltzmann’s constant, $T$ is the temperature, $A$ is the amplitude of the maximum ADC input signal and GBW is the gain stage amplifier Gain Bandwidth. This equation includes noise contributions from both the sampled thermal noise ($kT/C_{\text{sample}}$) and the amplifier thermal noise ($4kT/3g_{m1}T_{BW}$). In order to minimize the size of the sampling capacitor and therefore the size of the ADC input capacitance seen by the external driving circuit, the sampled thermal noise contribution is designed to dominate the $SNR_T$. Considering the gain stage sampled noise contribution with the assumption that the amplifier thermal noise contribution is negligible, the RMS noise is set 6dB below the quantization noise level, giving $SNR_T = 2^{N_{\text{tot}}/2}$. Substituting this $SNR_T$ value into (18), the size of the sampling capacitor in the gain stage is determined as

$$C_{\text{sample}} = \frac{16kT\cdot2^{2N_{\text{tot}}}}{(A)^2}.$$ \hspace{1cm} (19)

It is proposed that the gain stage utilises a two stage Miller compensated amplifier with a dominant pole, the step response gives a decaying exponential response (Laker and Sansen, 1994). From this simple model the Gain Bandwidth (GBW) may be determined as

$$GBW = \frac{(G+1)\ln(\frac{1}{\epsilon})}{T_{\text{amp}}2\pi},$$ \hspace{1cm} (20)

where $T_{\text{amp}}$ is the settling time of the gain stage and the settling error is $\epsilon$. The error at the gain stage output is required to be less than half an LSB (Least Significant Bit) of the second stage quantizer for the maximum signal at the gain stage input ($A/2$). The ratio of the output to the input gives the maximum settling error $\epsilon = 2^{-N_2}$.

The first stage current $I_{\text{bias1}}$ of the gain stage is set by the design choice to have the total amplifier thermal noise level at a lower $SNR_T$ than the sampled noise contribution. For the gain stage thermal noise contribution due to the amplifier, the RMS noise level is set 9dB below the quantisation noise of the complete ADC. Therefore substituting $SNR_T = 2^{N_{\text{tot}}/2}$ into (18), ignoring the sampled thermal noise contribution and solving for $g_{m1}$, the first stage amplifier current may then be expressed as

$$I_{\text{bias1}} = \frac{g_{m1}V_{\text{dsat1}}}{2} = \frac{8kT\cdot2^{2(N_{\text{tot}}+1.5)/\pi \cdot GBW \cdot V_{\text{dsat1}}}}{3\pi^2g},$$ \hspace{1cm} (21)

where $V_{\text{dsat1}}$ is the overdrive bias voltage for the differential pair transistors.

As the gain stage uses a two stage topology, it is recommended that the pole frequency is set to three times the GBW in order for the gain stage to be stable and have a phase margin greater than 60° (Laker and Sansen, 1994). The frequency of the output pole $\frac{g_{m2}}{2\pi C_L}$, which is dependent on the second stage transconductance and load capacitance, determines the second stage current consumption

$$I_{\text{bias2}} = \frac{g_{m2}V_{\text{dsat2}}}{2} = \frac{6\pi GBW C_L V_{\text{dsat2}}}{2},$$ \hspace{1cm} (22)

where $V_{\text{dsat2}}$ is the overdrive bias voltage for the output transistor, $g_{m2}$ is the transconductance of the amplifier second stage. The total load capacitance $C_L'$ is given by

$$C_L' = 2^{N_2-1}C_{\text{sar unit}} + \frac{C_{\text{sample}}}{g} + C_{\text{amp para}},$$ \hspace{1cm} (23)
where $2^{N_2-1}C_{\text{sar unit}}$ is the total input capacitance of the second stage quantizer, $C_{\text{sar unit}}$ is defined in (A2), $C_{\text{sample}}$ is defined in (19) and $C_{\text{amp para}}$ is the parasitic MOSFET (Metal Oxide Semiconductor Field Effect Transistor) device capacitance of the gain stage. The sum of the determined bias currents allows the power consumed by the gain stage to be approximated for a supply voltage of $V_{DD}$

$$P_{\text{amp}} = 2.2(I_{\text{bias1}} + I_{\text{bias2}})V_{DD}.$$  

(24)

The 2.2 factor is due to a factor of 2 for multiple current branches in a fully differential amplifier circuit and includes an allowance of 20% current overhead for additional bias circuitry in the amplifier.

(b) SAR ADC

Subtracting the DAC voltage from the input sample creates a residue voltage which is gained and passed to the second quantizer stage. The second quantizer stage is used for fine quantization (9 bits). Using a flash ADC as the second quantizer would increase the system power consumption as a flash ADC’s power doubles for bit each increase in resolution. A switched capacitor SAR offers a method of achieving the required fine quantization with low power consumption (Kull et al., 2013, Liu et al., 2010).

Previously Murmann (2012) analysed the power consumed by a SAR ADC by separating it into three components: the capacitor DAC ($P_{\text{cap dac}}$), the digital logic ($P_{\text{logic}}$); and the comparator. Murmann’s analysis looked at the minimum energy achievable by a SAR ADC, in contrast this analysis looks at the power consumed by a high speed or low latency SAR. In this paper the comparator power is further separated into the power consumed by the comparator latch ($P_{\text{comp latch}}$) and preamplifier stages ($P_{\text{comp preamp}}$), the SAR ADC power can be calculated as the sum of the powers of the SAR sub-blocks

$$P_{\text{SAR}} = P_{\text{cap dac}} + P_{\text{logic}} + P_{\text{comp latch}} + P_{\text{comp preamp}}.$$  

(25)

All of these blocks except the preamplifier are dynamic blocks so the energy $E$ required to switch them is of the general form

$$E = \frac{1}{2}CV_{DD}^2.$$  

(26)

This is the energy to charge or discharge a capacitor $C$ which can be a capacitor DAC or logic load and $V_{DD}$ is the supply voltage. The SAR implementation proposed in this work utilises a set and down switching scheme as described in Liu et al. (2010). This scheme has a DAC with $N_2 - 1$ binary weighted capacitors instead of the $N_2$ binary weighted capacitors as found in a conventional DAC implementation. The power consumption of the $N_2 - 1$ capacitor DAC is determined to be dependent on the number of capacitor switching events $(N_2 - 1)$ and the total SAR capacitance.

$$P_{\text{cap dac}} = \frac{1}{2}f_s \sum_{i=1}^{N_2-1} 2^{(N_2-1-i)}C_{\text{sar unit}}V_{\text{ref}}^2,$$  

(27)

where $V_{\text{ref}}$ is the reference voltage supplied to the ADC, and $f_s$ is the ADC sample rate. The size of SAR unit capacitance $C_{\text{sar unit}}$ is fixed by the minimum of feature size or thermal noise. This is further discussed in the appendix. The logic power $P_{\text{logic}}$ is based on a 100fF load capacitance which is switched every time a comparator bit is processed

$$P_{\text{logic}} = \frac{1}{2}C_{\text{logic}}V_{DD}^2 N_2 f_s.$$  

(28)

The comparator latch power is due to the charging and discharging of capacitors and is given by (Razavi, 2015) as

$$P_{\text{comp latch}} = (2C_p + C_{\text{out}})V_{DD}^2 N_2 f_s.$$  

(29)
where $C_p$ is an internal capacitor in the comparator and $C_{out}$ is the capacitance at the output of the comparator. The power in preamplifier stages of the comparator is determined by the overall gain $G_{preamp}$ in the preamplifier

$$P_{comp,preamp} = \frac{G_{preamp}}{G_{stage}} \cdot P_{stage}$$

where $G_{stage}$ is the gain per preamplifier stage and $P_{stage}$ is the power consumed by each stage. As detailed in the appendix (equation (A7)), the preamplifier gain is determined by the allocated SAR conversion time.

(c) Digital Prediction Filter

The quantized residue is passed to the filter stage. Unlike a conventional sigma delta ADC where the noise shaping filter is always implemented in the analog domain (Zrilic, 2006), it is possible to have a digital implementation of the filter in the predictive encoder approach. This is an advantage in low geometry CMOS process where the power consumption of high speed digital logic is low. Synthesis of an 8th order FIR filter with coefficients quantized to 9bits leads to a power consumption of 1mW in 90nm CMOS technology.

(d) Feedback DAC

The final block in the predictive encoder ADC is the switched capacitor feedback DAC where the digital output of the predictive filter is converted back to a voltage. The DAC in Figure 15 operates the same as in a pipeline ADC (Cline and Gray, 1996). The sampling capacitor and the feedback capacitor DAC share the same array of capacitors; the total sampling/DAC capacitor size is set by sampling noise requirements. The operation is as follows: firstly the input signal is connected to the capacitor DAC; sampling occurs when the switches open at the end of phase $\phi_1$; during $\phi_2$ the feedback DAC switches connect either $+V_{ref}$ or $-V_{ref}$ to the capacitors and settling of the DAC occurs.

The primary source of power consumption in the DAC is the energy to charge and discharge the capacitors (form given by equation (26)). The DAC capacitors sample the input signal level and the DAC input code tracks the input signal due to the prediction filter. The worst case charging condition occurs when the DAC is processing a sample with the input signal at the common mode level $V_{cm} = (V_{ref+} + V_{ref-})/2$, where $V_{ref+} = V_{cm} + V_{ref+}/2$ and $V_{ref-} = V_{cm} - V_{ref-}/2$. In this worst case half the capacitors must switch from to $V_{cm}$ to $V_{ref+}$ and the other half must switch to $V_{ref-}$. The energy required for this operation is

$$E_{DAC} = \frac{1}{2} C_{sample} (V_{ref+} - V_{cm})^2 + \frac{1}{2} C_{sample} (V_{cm} - V_{ref-})^2.$$  

The reference voltage $V_{ref}$ has a magnitude equal to $A$ the amplitude of the maximum differential signal at the ADC input

$$|V_{ref+} - V_{cm}| = |V_{ref-} - V_{cm}| = A/2.$$  

To convert energy to power the energy is multiplied by the sampling frequency $f_s$. A fully differential implementation is used so the power consumed due to the two DACs is

$$P_{DAC} = 2E_{DAC} f_s = \frac{C_{sample} A^2 f_s}{4}.$$  

A circuit realisation would require internal or external reference buffers to charge the DAC but estimation of the reference buffer power is not considered in this lower bound power estimation.
(e) System Timing & Power

The ADC consists of only one loop and is not pipelined so all the signal processing must occur within one clock period. Equation (34) below describes the block processing delays which must all take place in one clock period $T_{\text{clock}}$.

$$T_{\text{clock}} = T_{\text{amp}} + T_{\text{sar}} + T_{\text{fit}} + T_{\text{dac}}.$$  \hspace{1cm} (34)

Where $T_{\text{fit}}$ and $T_{\text{dac}}$ are the propagation times of the digital prediction filter and DAC respectively. The digital prediction filter logic is designed to have a minimal propagation time. The DAC capacitors settle at the same time as the gain stage so the DAC propagation delay included is only the delay due to the DAC logic and is therefore also short. Consequently the propagation delay of the loop is dominated by the sum of the gain stage settling and the SAR ADC propagation times. Therefore a key design consideration of the design of the system is the allocation of the fraction of the available times between the two blocks in order to minimise the total overall power consumption.

As a representative case study example, a 100MS/s sample rate $f_s$ and 10ns clock period $T_{\text{clock}}$ is chosen for the predictive encoder ADC. It is assumed that 2ns of the clock period is occupied by the digital prediction filter and DAC propagation times. The remaining 8ns may be therefore allocated between the gain stage and SAR ADC.

Based on the equations (24) and (27-30) the gain stage and SAR ADC power may be determined for given fractions of the remaining 8ns of the clock period. In this analysis the variable values used are detailed in Table 1, these values are based on expected parameters from a modern nanometre scale CMOS technology.

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>300K</td>
</tr>
<tr>
<td>$k$</td>
<td>$1.38 \times 10^{-23}$ m$^2$ kg s$^{-2}$ K$^{-1}$</td>
</tr>
<tr>
<td>$N_{\text{tot}}$</td>
<td>12 bits</td>
</tr>
<tr>
<td>$N_1$</td>
<td>3 bits</td>
</tr>
<tr>
<td>$N_2$</td>
<td>9 bits</td>
</tr>
<tr>
<td>$f_s$</td>
<td>100MHz</td>
</tr>
<tr>
<td>$G$</td>
<td>8</td>
</tr>
<tr>
<td>$V_{\text{dd}}$</td>
<td>1V</td>
</tr>
<tr>
<td>$V_A$</td>
<td>0.8V</td>
</tr>
<tr>
<td>$V_{\text{sat}}$</td>
<td>0.15V</td>
</tr>
<tr>
<td>$V_{\text{ref}}$</td>
<td>0.8V</td>
</tr>
<tr>
<td>$C_{\text{min}}$</td>
<td>1fF</td>
</tr>
<tr>
<td>$C_{\text{amp,para}}$</td>
<td>200fF</td>
</tr>
<tr>
<td>$C_{\text{logic}}$</td>
<td>100fF</td>
</tr>
<tr>
<td>$C_p$</td>
<td>20fF</td>
</tr>
<tr>
<td>$C_{\text{out}}$</td>
<td>50fF</td>
</tr>
<tr>
<td>$R_{\text{sw}}$</td>
<td>1k$\Omega$</td>
</tr>
<tr>
<td>$C_{\text{stage}}$</td>
<td>2</td>
</tr>
<tr>
<td>$T_{\text{conv,av}}$</td>
<td>5</td>
</tr>
<tr>
<td>$P_{\text{stage}}$</td>
<td>1mW</td>
</tr>
<tr>
<td>$t_0$</td>
<td>10.5ps*</td>
</tr>
<tr>
<td>$\tau_{\text{cmp}}$</td>
<td>4.65ps*</td>
</tr>
<tr>
<td>$T_{\text{logic}}$</td>
<td>20ps</td>
</tr>
</tbody>
</table>

Table 1. Summary of variable values. *measured by Kull et al (2013).
The results are plotted in Figure 16. On the left side region of the figure the allocated SAR conversion time is maximised and power consumption is consequently dominated by the gain stage. This occurs as the gain stage bandwidth increases in order to achieve faster settling. On the right side of the Figure 16, the gain stage settling time is maximised, its power consumption is reduced and the SAR power consumption now dominates total power. The reduction in the SAR propagation times can only be achieved by adding power consuming pre-amplification gain stages to the comparator. The logarithmic relationship between preamplifier gain and comparator settling time as detailed in equation (A5), causes the rapid increase in the SAR ADC power consumption as additional gain is added as settling time is reduced. It is observed from figure 16 that the optimal system power consumption is found when 6ns are allocated to the amplifier and 2ns are allocated to the SAR ADC. The estimated power consumption for the predictive ADC system blocks using optimal allocation of the propagation times for the gain stage and SAR ADC are detailed in Figure 17. It is noted that this figure provides a lower bound for power consumption for the predictive encoder based architecture. The trade off in power consumption based on allocation of time between the gain stage settling time and the SAR ADC operation is a key consideration for a transistor level implementation of the predictive encoder architecture.

8. Conclusions.
This paper has demonstrated that feedback predictive based encoder ADCs have limited input frequency ranges for given input amplitudes. The input bandwidth may be maximized by appropriate selection of filter order, choice of system gain and loop quantizer resolution. The predictive based encoder ADC structure has a similar no-overload condition to the sigma delta modulator ADC, but is dominated by prediction error rather than quantization noise. The overload condition is both input amplitude and input frequency dependent. Limit cycle oscillation may occur with the use of higher order filters in predictive encoder ADC configurations, however with appropriate selection of prediction filter order the phase margin of the resulting integrator should be sufficient to avoid limit cycle oscillations. A switched capacitor based realization of the predictive encoder is proposed. Optimisation of power consumption based the allocation of time between the gain stage and the SAR ADC operation is investigated and the lower bound of power consumption for this architecture is estimated.

References.


Appendix

The purpose of this appendix is the determination of two results that are necessary for calculation of the power consumption of the SAR ADC as discussed in Section 7(b). Firstly the settling time for the SAR DAC capacitor array based on the minimum size of the SAR capacitor is determined. Secondly an expression for comparator settling time is determined. The required comparator preamplifier gain to achieve settling within a given time specification may be used to estimate the SAR comparator power consumption.

(a) SAR DAC capacitor array settling time.

The minimum SAR DAC capacitance $C_{\text{sar,min}}$ is calculated to be the total amount of DAC capacitance required to keep the thermal noise 12dB below the quantisation noise floor ($SNR_F = 2^{N_2}/2^{-2}$) of the SAR ADC

\[ C_{\text{sar,min}} = 4kT \cdot 2^{2(N_2+2)} \cdot \frac{1}{f^2}. \]  

(A1)

The SAR unit capacitor $C_{\text{sar,unit}}$ is calculated to be the largest of either $C_{\text{sar,min}}$ divided by $2^{N_2-1}$ or the smallest unit capacitor achievable in a given process technology $C_{\text{min}}$

\[ C_{\text{sar,unit}} = \max \left( \left( 128kT \cdot 2^{N_2} \cdot \frac{1}{f^2} \right), C_{\text{min}} \right). \]  

(A2)

It is assumed that the SAR capacitor DAC can be trimmed to achieve the desired matching and the minimum size is not limited by device mismatch. The analysis of the SAR capacitor DAC settling time assumes that the DAC switches are small so they can be easily driven by logic and don’t require multiple buffers which would increase the propagation delay. The switch resistance $R_{\text{sw}}$ and the capacitor require a certain number of RC time constants to settle to the voltage level of half a SAR LSB. For the MSB (Most Significant Bit) half of the capacitance is switched and the capacitance reduces by two for each subsequent bit. As a simplification it is assumed that only the switched caps have to settle, in reality the entire capacitor bank has to settle but for a smaller voltage step. The total SAR DAC capacitor array settling time $T_{\text{Cap.settle}}$ can be written as a summation

\[ T_{\text{Cap.settle}} = \sum_{i=1}^{N_2-1} R_{\text{sw}} 2^{N_2-i-1} C_{\text{sar,unit}} \ln 2^{N_2+1-i}. \]  

(A3)

(b) Comparator settling time & preamplifier gain.

Consider a time $T_{\text{SAR}}$ is allocated for the SAR conversion, during which the $N_2$ bit SAR ADC must complete $N_2$ comparator decisions, and $N_2-1$ SAR capacitor DAC switching and settling events. The complete SAR conversion time may therefore be expressed as

\[ T_{\text{SAR}} = T_{\text{Cap.settle}} + [T_{\text{comp}} + T_{\text{logic}}] N_2 \]  

(A4)

where $T_{\text{Cap.settle}}$ is determined by (A3), $T_{\text{comp}}$ is the time for a single comparator decision and $T_{\text{logic}}$ is the digital logic propagation delay for processing a single comparator decision. The logic time $T_{\text{logic}}$ is the digital propagation time which is determined by the CMOS process technology chosen and is assigned a fixed value in this analysis.

The comparator is assumed to be a dynamic type comparator or Strong-Arm latch as described in Razavi, (2015). Use of this structure allows the option of including wideband preamplification stages at the input to speed up the overall comparator decision time. The comparator preamplifier gain increases the latch input voltage. This increased voltage means less latch time constants are required for the comparator to correctly resolve the input voltage to the required precision of the SAR DAC. The comparator preamp stages are wideband.
stages which provide small linear gain $G_{\text{preamp}}$. A result for the comparator decision time is derived in Kull et al., (2013),

$$T_{\text{comp}} = t_0 + \tau_{\text{cmp}} |\ln(V_{icmp} \cdot G_{\text{preamp}})|,$$

$$= t_0 + \tau_{\text{cmp}} (|\ln(V_{icmp})| - \ln(G_{\text{preamp}})), \quad \{ V_{icmp} < 1, G_{\text{preamp}} > 1 \}$$

(A5)

where $t_0$ is the fixed time of the integrating gain stage of the dynamic comparator, $\tau_{\text{cmp}}$ is the latch time constant, $V_{icmp}$ is the comparator input voltage and $G_{\text{preamp}}$ is the preamplifier gain. Values for the constants $t_0$ and $\tau_{\text{cmp}}$ measured by Kull et al., (2013) for a modern nanometre CMOS technology are also used in this analysis. In order to further simplify the expression (A5) and remove the unknown varying quantity $V_{icmp}$, the result of Waters et al. (2015) is used. It is demonstrated that for a 10bit asynchronous (self-timed) SAR a total conversion time of less than $50 \tau_{\text{cmp}}$ achieves a signal to metastability to error ratio (SMR) of 70dB. This result sets an upper bound on the average number of time latch constants required by the comparator without pre-amplification ($G_{\text{preamp}} = 1$) for each conversion. For each comparator bit decision this is equal to an average number of latch time constants $T_{\text{latch,av}} = \tau_{\text{cmp}} |\ln(V_{icmp})| = 5\tau_{\text{cmp}}$. Substituting for $\tau_{\text{cmp}} \ln(V_{icmp})$ in (A5), leads to the expression

$$T_{\text{comp}} = t_0 + (T_{\text{latch,av}} - \tau_{\text{cmp}} \ln(G_{\text{preamp}})), \quad 1 < G_{\text{preamp}} < e^{T_{\text{latch,av}}},$$

(A6)

Adding preamplifier gain now acts to reduce the overall comparator decision time. It is assumed that the preamp stages have minimal delay. Re-arranging and substituting (A4) gives the expression for the required preamplifier gain to achieve conversion within a specified SAR conversion time $T_{\text{SAR}}$

$$G_{\text{preamp}} = e^{(T_{\text{cap,settle}} - T_{\text{SAR}} + (T_{\text{logic}} + t_0 + T_{\text{latch,av}})N_2)/(\tau_{\text{cmp}} N_2)}.$$

(A7)
Figure 1. Equivalent quantizing systems with both providing input tracking. (a) 2 stage Pipeline ADC. (b) Predictive encoder ADC.
Figure 2. Encoders: (a) Predictive, (b) Sigma Delta Modulator. ADCs: (c) Feedback Predictive, (d) Sigma Delta.
Figure 3. Relationship between Predictive Encoder and Sigma delta modulator, shown for a 2nd order cascade of integrators structure.
Figure 4. (a) Prediction Filter $H_p$ (form (5)) magnitude and phase response, $n = 11$, $\alpha = 0.4$. (b) Associated high pass filter magnitude and phase response ($H_{hp} = 1 - H_p$).

Figure 5. Predictive Encoder discrete time equivalent
Figure 6. SINAD, Noise and prediction error power variation, $n = 7$, $A = 0.9$, $\alpha = 0.4$, $G = 8$, $N_2 = 6$.

Figure 7. Normalized input frequency vs predictive filter order (filter form (5)), $\alpha = 0.4$, for varying values of input amplitude $A$, ($G = 8$, $N_2 = 9$).
Figure 8. Normalized input frequency vs System Resolution (Increasing Gain) for filters of form (5), $\alpha = 0.4$, for varying values of filter order $n$. $N_2 = 9$.

Figure 9. Predictive encoder ADC open loop gain and phase, predictive filter form (5), orders $n = 7$ to $13$, $\alpha = 0.4$
Figure 10. Forward path magnitude at expected oscillation frequency versus Gain/Saturation block input amplitude for various prediction filter orders $n$.

Figure 11. Theoretical & Measured Oscillation Frequency versus Oscillation Amplitude at input to Gain/Saturation block.
Figure 12. Harmonic Mode operation, predictive filter (form (5)), $n = 7$, $A = 0.5$, $\alpha = 0.4$. $f_{\text{in}} = 0.7060 \pi \text{ rad/sample}$. Aliased 1st harmonic peak observed at $\pi 0.12 \text{ rad/sample}$.

Figure 13. Normalized input frequency vs predictive filter order (form (5)): Theory (14) vs Simulink model results. ($A = 0.9$, $N_2 = 9$, $G = 8$, $\alpha = 0.4$).
Figure 14. Comparison of SINAD sweeps for two predictive encoder 12bit ADCs with different quantizer sizes and gains for the same stable filter order (n = 8).
Figure 15. Discrete time circuit level realisation of the predictive encoder based ADC.
Figure 16. Power Tradeoff of Total Gain Stage and SAR ADC Power versus allocated time

Figure 17. Estimated power consumption for each of the predictive encoder based ADC sub-blocks in Figure 15.