ANALOG-TO-DIGITAL CONVERSION

Analog-to-digital (A/D) converters (ADCs) constitute the key interface function between analog signals in the physical world and digital signal processing systems. The importance of integrated ADCs has grown enormously in recent years, in line with the increasing importance of mixed analog-digital VLSI systems. Indeed, with the powerful digital signal processing engines available today the fidelity and accuracy of digitally processed analog signals is fundamentally limited by

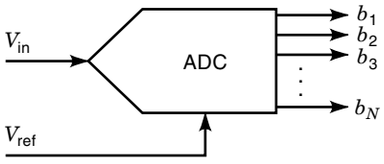


Figure 1. Symbol of an ADC.

the performance of ADCs rather than by any meaningful digital circuitry limitation. Moreover, given the continuing need to achieve higher integration functionality at minimum cost and with minimum energy consumption, the design of ADCs is becoming increasingly tailor-made to specific applications rather than, as not long ago, being considered as general purpose components (1). As a result, a wide variety of ADC architectures and circuit design techniques has emerged in recent years, ranging from low to high conversion frequency as well as from high to low conversion resolution. The most relevant of such conversion architectures and circuit design techniques are described in this article.

FUNDAMENTAL ASPECTS OF A/D CONVERTERS

Conversion Between Analog and Digital Samples

Ideal Conversion Characteristic. An ADC is usually represented by the symbol indicated in Fig. 1, where the number of bits N of the output digital word indicates the conversion resolution. The conversion characteristic of an ADC is obtained by determining the transition voltages $(V_T)_n$ of the input analog signal that produce a change of the output digital code. The time interval between two consecutive input samples indicates the sampling period T_s. The conversion frequency is $F_s = 1/T_s$.

Figure 2 illustrates a portion of the ideal conversion characteristic of an ADC. The analog input transition voltages are defined by

$$V_{T_n} = \frac{V_{ref}}{2^N} n \tag{1}$$

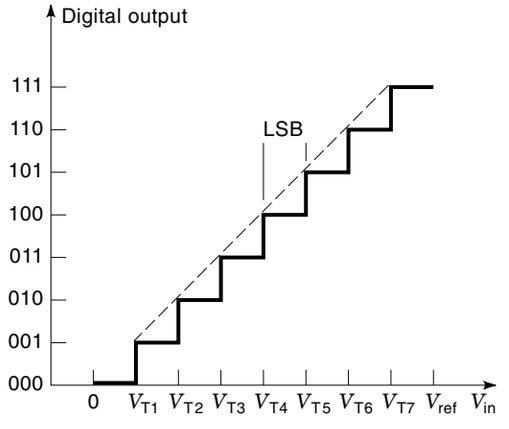


Figure 2. Ideal conversion characteristic of an ADC.

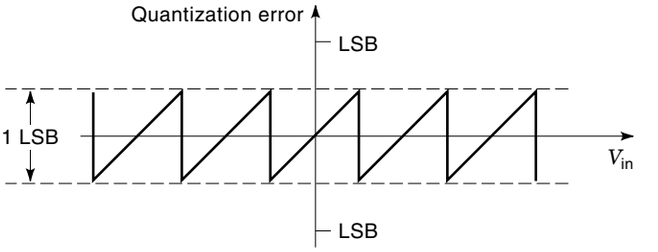


Figure 3. Quantization error produced by an ADC when the input signal is a ramp.

where V_{ref} is the reference voltage of the converter and

$$n = \sum_{i=1}^{N} 2^{i-1} b_i \tag{2}$$

represents the output digital code. The difference between two consecutive quantization levels is usually designated the least significant bit (LSB) of the converter. From Eq. (1) this is given by

$$\text{LSB} = \frac{V_{ref}}{2^N} \tag{3}$$

Quantization Noise. Even assuming that the conversion characteristic of an ADC is ideal, the quantization process produces a quantization error signal that is a function of the conversion resolution. This is often called the *quantization noise* of the converter (2). Figure 3 illustrates the shape of such a quantization error signal when the input signal is a ramp. Assuming that the maximum quantization error is uniformly distributed between $-$LSB/2 and $+$LSB/2, then it can be shown that the total quantization noise power can be expressed as

$$P_Q = \frac{\text{LSB}^2}{12} \tag{4}$$

In most practical cases the quantization noise power is uniformly distributed in the Nyquist frequency band from dc to $+F_s/2$, as depicted in Fig. 4, such that the corresponding power spectral density $S_Q^2(f)$ is expressed as

$$S_Q^2(f) = \frac{\text{LSB}^2}{12}\left(\frac{2}{F_s}\right) \tag{5}$$

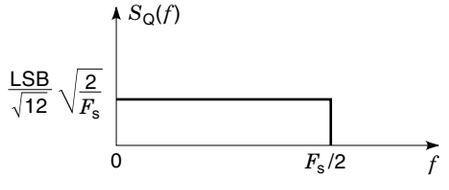


Figure 4. Uniform power spectral density of the quantization noise.

For any given frequency band within the Nyquist band, that is, from dc to $f_c < F_s/2$, the total noise power is given by

$$\int_0^{f_c} \left(\frac{\text{LSB}}{\sqrt{12}} \frac{1}{\sqrt{F_s/2}} \right)^2 df = \frac{\text{LSB}^2}{12} \left(\frac{2f_c}{F_s} \right) \tag{6}$$

Equation (6) indicates two possible solutions for reducing the quantization noise of a converted signal at a given frequency. One solution is to reduce the converter LSB by increasing its resolution N, whereas the alternative solution corresponds to having a conversion frequency much higher than the signal being converted. This is called oversampling and the factor of oversampling the signal would be defined by

$$\text{OSR} = \frac{(F_s/2)}{f_c} \tag{7}$$

Increasing the resolution of an ADC gives a reduction of the noise power of approximately 6 dB per added bit of resolution. As this would be guaranteed over the full Nyquist range, such type of converters are called *Nyquist converters.* In the so-called *oversampling converters,* a further reduction of the quantization noise power is obtained by oversampling the signal with respect to the Nyquist frequency; it yields an additional gain of approximately 3 dB per octave of oversampling.

Dynamic Range of a Quantized Signal. The dynamic range of a quantized signal is commonly determined with respect to a sinusoidal waveform between 0 and V_{ref}, whose power is $V_{\text{ref}}^2/8$. Thus, from Eq. (4), the resulting signal-to-noise (SNR) ratio of the quantized signal will be given by

$$\text{SNR(dB)} = 10 \log \left(\frac{P_{\text{in}}}{P_Q} \right) = 10 \log \left(\frac{V_{\text{ref}}^2/8}{\text{LSB}^2/12} \right) \tag{8}$$

From Eq. (3) this results in

$$\text{SNR(dB)} = 10 \log \left(\frac{3}{2} 2^{2N} \right) = (6.02N + 1.76)\,\text{dB} \tag{9}$$

For example, the conversion of an analog sine waveform using a 10-bit ADC will lead to a digital sine waveform with an SNR of approximately 61.8 dB.

Conversion Codes. When an ADC operates with only positive (or negative) signals its output is normally a natural binary code. There are many cases, however, when the ADC has to operate with both positive and negative signals, thus resulting in the need to produce both positive and negative output digital codes. Two of the most popular of such codes are the sign magnitude code and the 2's complement code.

In the sign magnitude code, negative and positive digital numbers are represented by the same code except for the most significant bit (MSB): for positive codes MSB = 0, whereas for negative codes MSB = 1. For example, the sign-magnitude representation for 7 is 0111, whereas for −7 it is 1111.

In the 2's complement code, positive codes are represented as natural binary codes with MSB = 0. The negative codes are obtained from the equivalent positive number by first complementing all the bits and then adding 1 LSB. For example, for the representation of −7 we obtain first 0111 (corresponding to the positive number +7), then we obtain 1000 (corresponding to the complement of all bits of 0111) and finally we obtain code 1001 (adding 1 LSB to 1000). An important advantage of the 2's complement code is that the addition of both positive and negative numbers is made by straightforward addition and no extra computations are needed. For example, 0111 (number +7) added to 1001 (number −7) gives 0000 (of course the carry bit is neglected).

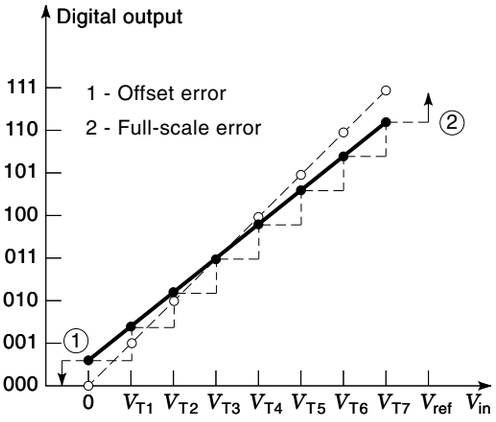


Figure 5. Illustrating offset, full scale and gain errors of an ADC.

ADC Limitations and Performance Parameters

Offset, Full Scale and Gain Errors. Figure 5 represents both the ideal and actual conversion characteristics of a 3-bit ADC (3). The former is represented by the broken line that unites the white dots corresponding to the ideal input transition voltages between all digital codes in the absence of conversion errors, whereas the latter is represented by the full line that unites all black dots corresponding to the actual analog transition voltages between all digital codes. On the lower extreme of the conversion characteristic, the *offset error* corresponding to ① represents the difference between the actual and ideal input voltage that produces a transition of one LSB from the output digital code 0. On the upper extreme of the conversion characteristic, the *full-scale error* corresponding to ② represents the difference between the actual and ideal input voltages that produce the transition of the output to the full scale digital code. The *gain error* of the ADC corresponds to the difference between the ideal slope of the conversion characteristic and its actual value.

Differential and Integral Nonlinearity. The *differential nonlinearity* (DNL) in the ADC expresses the deviation of the difference between the actual input transition voltages $V_a(n)$ and $V_a(n + 1)$ that produce any two consecutive output digital codes from the difference voltage corresponding to one LSB. Such LSB must be determined through a linear characteristic with the same offset and full-scale errors. This gives a measure of the linearity of operation for small incremental signals and can be expressed by

$$\text{DNL}(n) = \frac{V_a(n+1) - V_a(n)}{\text{LSB}} - 1 \tag{10}$$

for any digital code between $n = 1$ and $n = 2^N - 2$. The normal specification for the DNL of the ADC is $< \pm(1/2)$ LSB. It should also be mentioned that when the DNL is larger than 1 LSB a phenomenon of *nonmonotonicity* occurs whereby the output digital code decreases with an increase of the analog input signal.

By contrast with the DNL, the *integral nonlinearity* (INL) in the ADC gives a measure of the linearity of operation for large signals, as it expresses the deviation between the actual input transition voltages obtained for a given digital code and the corresponding ideal transition voltages assuming a linear characteristic with the same offset and full-scale errors. For any digital code from $n = 1$ to $n = 2^N - 1$, this can be defined as

$$\mathrm{INL}(n) = \frac{V_a(n) - V_a(1)}{\mathrm{LSB}} - (n - 1) \tag{11}$$

and relates to the DNL through the equation

$$\mathrm{INL}(n) = \sum_{i=1}^{n-1} \mathrm{DNL}(i) \tag{12}$$

The normal specification for the INL of an ADC is also $< \pm(1/2)$ LSB. Figure 6 illustrates the preceding definitions of the DNL and INL of an ADC.

Sampling Jitter. An important dynamic error that affects the performance of an ADC, especially for high-frequency operation, is due to the timing uncertainty of the sampling instants that produce the input analog samples for conversion (4). Because of this error, commonly called *sampling jitter,* the actual sequence of analog samples is not equally spaced in time but rather varies with respect to the nominal ideal sampling instants, as illustrated in Fig. 7(a). For a fixed conversion frequency, it is easily seen that the resulting signal error, represented in Fig. 7(b), increases as the frequency of the input analog signal increases. It can be shown that in order to obtain an error signal smaller than the LSB of the converter the sampling jitter must be such that

$$\Delta t_s < \frac{1}{\pi 2^{N+1} f} \tag{13}$$

where f is the frequency of the input signal. For example, for a 10-bit ADC at 1 MHz this corresponds to a maximum timing uncertainty of 310 ps.

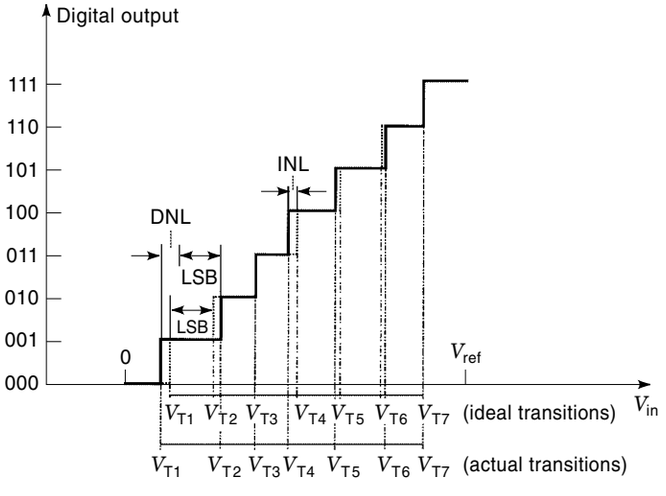


Figure 6. Integral nonlinearity (INL) and differential nonlinearity (DNL) of an ADC.

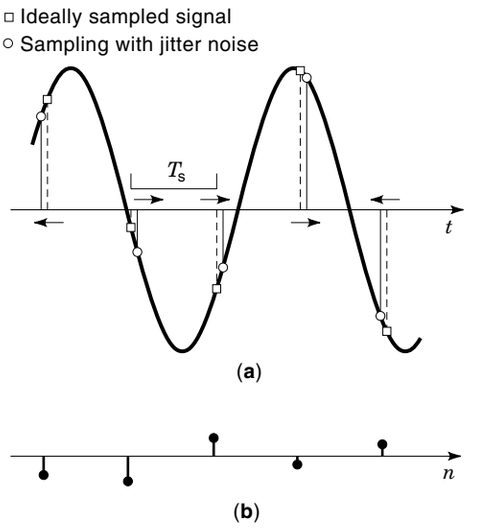


Figure 7. Illustrating the effect of sampling jitter. (a) Sampled sine waveform. (b) Jitter related error signal.

Effective Number of Bits. A global parameter that summarizes the performance behavior of an ADC is the effective number of bits (ENOB) expressed as

$$\mathrm{ENOB} = \frac{\mathrm{SINAD_{dB}} - 1.76}{6.02} \tag{14}$$

where the parameter $\mathrm{SINAD_{dB}}$ (signal-to-noise-plus-distortion ratio) represents the combined effect of quantization and harmonic distortion due to the various sources of nonlinearity of the converter. An example of the graphical evolution of the ENOB is given in Fig. 8, for an ADC with $F_s = 40$ MHz conversion frequency, showing that the effective number of bits of the converter decreases as the signal-to-noise-plus-distortion ratio decreases for higher frequencies of input analog signal.

NYQUIST ANALOG-TO-DIGITAL CONVERTERS

From the standpoint of the conversion frequency, we can classify ADCs into three main groups. Serial converters are the slowest of all because each bit of the output digital word is usually determined in a number of clock cycles that rises proportionally to the equivalent bit weight. For example, for the

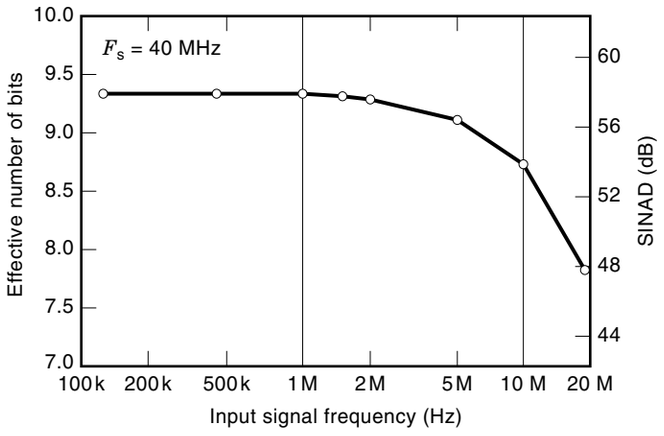


Figure 8. Evolution of the effective number of bits of an ADC as a function of the frequency of the input signal.

LSB this may take only one clock cycle, but for the most significant bit this may take as long as 2^N clock cycles. By contrast with their low speed, serial processing converters offer the highest resolution of all the converters described in this section. They are therefore particularly useful to interface slowly varying signals for very high precision signal processing as required, for example, in instrumentation and biomedical applications.

Following the serial converters in the speed scale is the group of algorithmic ADCs. Here, conversion takes place in a number of clock cycles that is typically proportional to the conversion resolution. Hence, depending on the actual frequency limitations of the constituting building blocks, conversion frequencies of a few megahertz can be achieved without too much difficulty in modern complementary metal oxide semiconductor (CMOS) technology. Algorithmic ADCs constitute a versatile group of converters that can meet a wide range of specifications in terms of the conversion frequency as well as resolution. A particular type of ADC based on the successive approximation algorithm constitutes an industry workhorse for a large variety of applications.

The fastest group of converters is based on parallel processing techniques that allow a full conversion period to be performed in only one or, at most, a few (e.g., 2) clock cycles. One-clock cycle (flash) ADCs in modern CMOS technology can reach conversion frequencies above 100 MHz but their resolution is limited to no more than 8-bits. An increasingly popular group of parallel ADCs is based on pipeline architectures that can perform a full conversion period in only two clock cycles, although there is an initial latency period that depends on the total number of stages of the pipe. Pipeline converters with optimized silicon and power dissipation are in great demand for video processing applications for the consumer market.

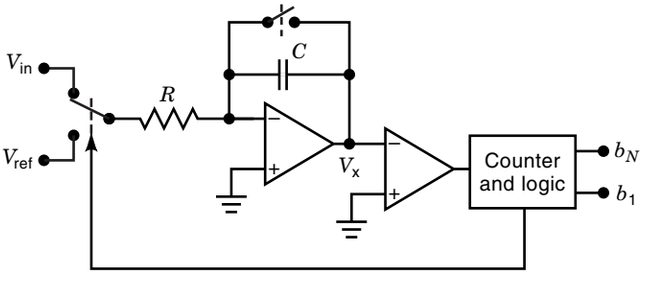
Serial Processing ADCs

Double-Ramp Integration. One of the best known architectures for serial processing analog-digital conversion is illustrated in Fig. 9(a), comprising an active-RC integrator with multiplexed input, a comparator and a digital counter and control logic. Figure 9(b) illustrates the time-domain waveforms typically produced in the converter during a complete conversion cycle. At the beginning of the conversion cycle the capacitor around the feedback loop of the opamp is reset to zero. Then, the integrator input is connected to the input signal terminal to generate the first integration ramp that is carried out during a fixed time interval T_1 controlled by the digital counter. Thus, for a clock period of T_s the opamp output voltage reached after the first integration ramp is

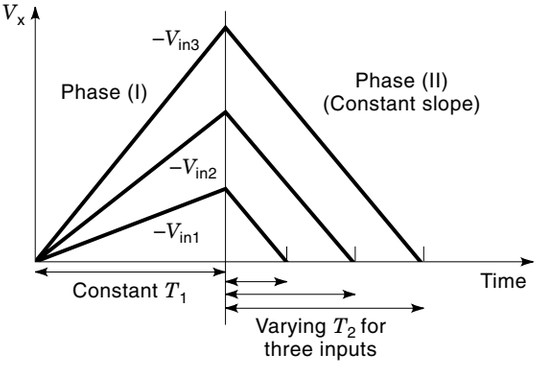
$$V_x = \frac{N_1 T_s}{RC}(-V_{in}) \tag{15}$$

where N_1 is the fixed number of clock cycles counted during T_1. Next, the input terminal is switched to the reference voltage terminal and a second integration ramp is generated whose slope is the negative of the first integration ramp. This is carried out until the comparator detects that the output voltage of the amplifier crosses zero and thereby stops the digital counter. The duration of this second integration ramp, $T_2 = N_2 T_s$, varies as the voltage at the output of the amplifier at the end of the first integration ramp is larger or smaller. This is determined from

$$0 = V_x + \frac{N_2 T_s}{RC} V_{ref} \tag{16}$$



(**a**)



(**b**)

Figure 9. Serial processing ADC with double-ramp integration. (a) Typical circuit realization. (b) Waveforms.

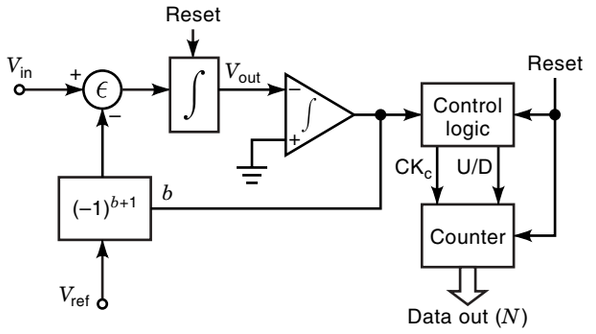


Figure 10. Conceptual block diagram of an incremental ADC.

and which, combined with Eq. (15), yields

$$V_{\text{in}} = \frac{N_2}{N_1} V_{\text{ref}} \tag{17}$$

Hence, the variable time interval of the second integration ramp (expressed in terms of the counted number of clock cycles) is a measure of the value of the analog input signal integrated during the first integration ramp.

Because the conversion function Eq. (17) is independent of the passive circuit elements it allows very high conversion resolutions to be achieved, although at rather low conversion frequencies. Such high-resolution and low-speed conversion characteristics render double-ramp ADCs particularly suitable for applications in telemetry, instrumentation, and measurements.

Incremental Converters. Another type of serial processing ADC is based on the so-called incremental converter (5,6). Its conceptual block diagram is illustrated in Fig. 10 where, as before, the processing core of the converter is formed by an integrator and a comparator. The control logic and digital counter provide the additional digital processing functions required for conversion. The main difference with respect to the previous double-ramp converter lies in the fact that the integration variable is now formed by the difference between the sampled input signal and an analog signal determined as a function of the output voltage of the comparator.

A switched-capacitor implementation of the processing core of an incremental ADC is shown in Fig. 11(a). It is assumed that during a full conversion cycle the sampled input signal of the converter is held constant. For simplicity of the following discussions it is further assumed that the input sampled and held signal is positive. At the beginning of the conversion cycle, the integrating capacitor is reset by closing switch SR. In each subsequent conversion step switches S1, S3, S4, and S5 are controlled in such a way as to produce an incremental variation of $\Delta V_x = -(C_i/C_f)V_{in}$ at the output voltage of the opamp. Then, whenever V_x reaches zero, a fraction of the reference voltage, actually $(C_i/C_f)V_{ref}$, is subtracted from the output using switches S2, S3, S4 and S5, and the digital counter is advanced by 1. The typical waveform observed at the output of the opamp is depicted in Fig. 11(b). At the end of 2^N conversion steps the output voltage becomes

$$V_x = 2^N V_{\text{in}} \frac{C_i}{C_f} - n \times V_{\text{ref}} \frac{C_i}{C_f} \tag{18}$$

where $-V_{\text{ref}}(C_i/C_f) \le V_x \le 0$ and n is the output of the counter (digital integrator). This implies

$$n = 2^N \left(\frac{V_{\text{in}}}{V_{\text{ref}}} \right) + \epsilon \tag{19}$$

where the residual error is $0 \le \epsilon \le 1$. Thus, n is an N-bit digital representation of the sampled input voltage, with a quantization error $\epsilon \le 1$ LSB.

As in the double-ramp integrator, the conversion characteristic expressed by Eq. (19) is independent of the capacitance ratio and, hence, resolutions of the order of 16-bits and above can also be achieved without too much difficulty. Of course this will be possible only for very low frequency of conversion due to the serial processing nature of the converter.

Algorithmic A/D Converters

Next in the scale of conversion frequency is the group of algorithmic ADCs where the number of clock cycles needed for conversion is directly, rather than exponentially, proportional to the conversion resolution. Two types of algorithmic ADCs will be described. The first is based on the classical successive approximation algorithm and constitutes the industry workhorse for a large variety of applications. The second is based on a cyclic divide-by-two and subtract algorithm whose circuit implementations can be made more compact than in the case of the successive approximation ADC.

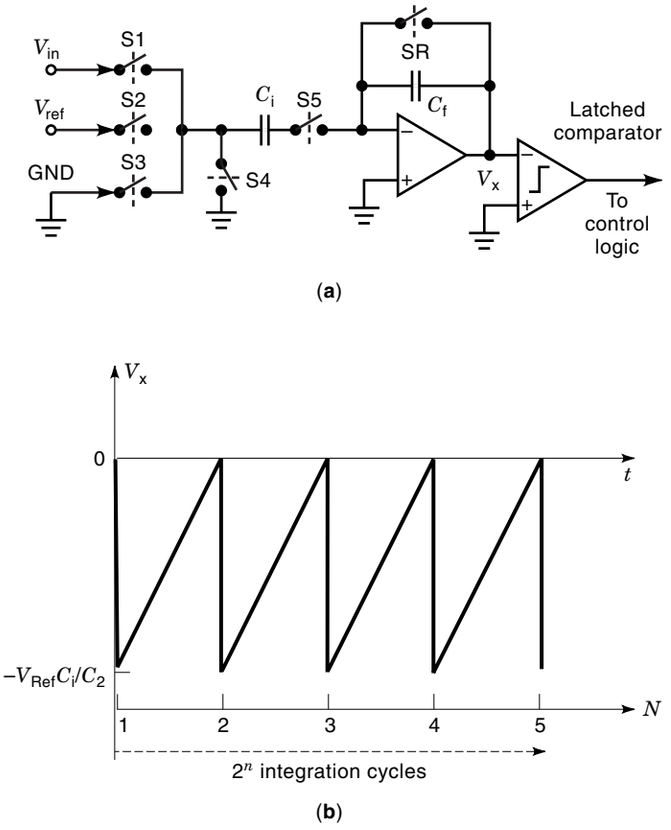


Figure 11. Switched-capacitor incremental ADC. (a) Circuit diagram and (b) Typical output waveform.

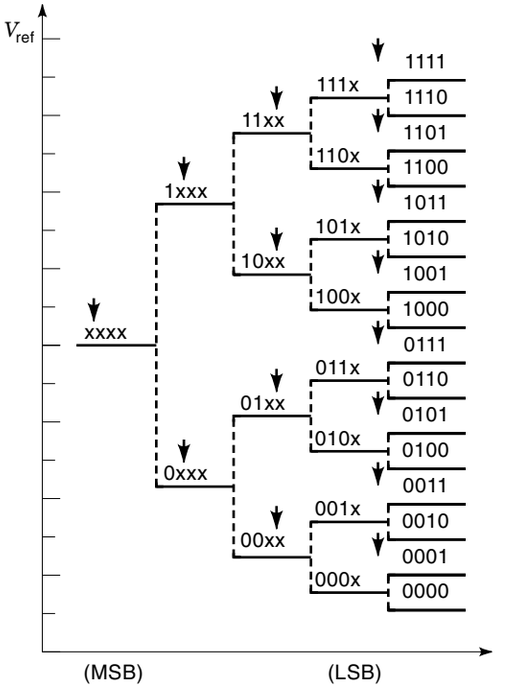


Figure 12. Illustrating the successive approximation algorithm for analog-to-digital conversion.

Successive Approximation ADC. The successive approximation ADC is based on the algorithm schematically represented in Fig. 12, for an example of 4-bits conversion. The equivalent analog weights of the digital bits are $V_{ref}/2$, for bit b_4 (the MSB), $V_{ref}/4$ for bit b_3, $V_{ref}/8$ for bit b_2, and finally $V_{ref}/16$ for the LSB (bit b_1). The execution of the algorithm is done from the MSB to the LSB during a number of clock cycles equal to the resolution of conversion. At the beginning of the conversion, the input signal is compared with the analog weight of the MSB and the result is either $b_4 = 1$ or $b_4 = 0$, depending on whether the signal is above or below $V_{ref}/2$. If the result is '1,' then the analog weight of the MSB is subtracted from the input signal; otherwise the input signal remains unchanged. In the next phase, the available analog signal is compared with the analog weight of bit b_3 and the result is either $b_3 = 1$ or $b_3 = 0$, depending on whether the signal is above or below $V_{ref}/4$. If the result is '1,' then the analog weight of bit b_3 is subtracted from the analog signal; otherwise the signal remains unchanged. Similar operations are carried-out in two more steps until the LSB is resolved.

Figure 13 represents a possible conceptual block diagram for the implementation of the forementioned successive approximation algorithm. The input sample-and-hold block provides the analog sampled signal for conversion. The digital-to-analog converter generates the equivalent analog values of the digital words containing the bits that are sequentially resolved, from the MSB to the LSB. The comparator acts as the decision element that indicates to the digital successive approximation register how the input analog signal is being approximated by the reconstructed analog values.

There are various possible circuit solutions for the implementation of the block diagram in Fig. 13, some employing digital-to-analog converters with resistive division and some

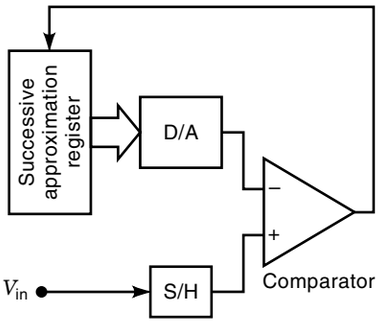


Figure 13. Block diagram of a successive approximation ADC.

employing digital-to-analog converters with capacitive division (1). The latter is one of the most popular forms of implementation, especially for CMOS technology, as it uses a capacitor array that can also provide the additional functions of sample-and-hold and subtraction needed in the conversion algorithm.

Figure 14 illustrates the implementation of a successive approximation ADC using an array of switched binary-weighted capacitors. During the execution of one conversion cycle the circuit is sequentially reconfigured as illustrated in Fig. 15. First, as seen in Fig. 15(a), the opamp is connected in a unity-gain feedback configuration while the capacitors are connected to the input terminal. Because of the virtual ground created at the negative terminal of the opamp the input voltage signal is sampled onto the top plate of the capacitors. In the next phase, shown in Fig. 15(b), the most significant capacitor is connected to the reference voltage and, as a result, a process of charge redistribution between this and the remaining capacitors takes place, yielding a new voltage expressed by

$$V_x = -V_{in} + \frac{V_{ref}}{2} \tag{20}$$

at the negative terminal of the open loop opamp. If this is negative, then the output voltage indicates the equivalent logic value '1'; otherwise it indicates the equivalent logic value '0.' In the former case, the MSB is set to '1' and the corresponding branch remains connected to the reference voltage.

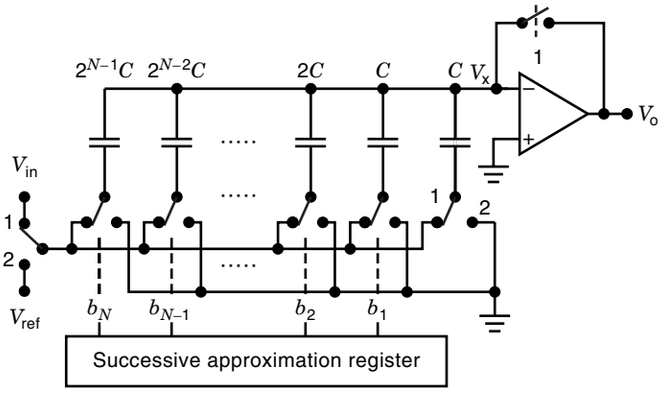


Figure 14. Switched-capacitor implementation of a successive approximation ADC.

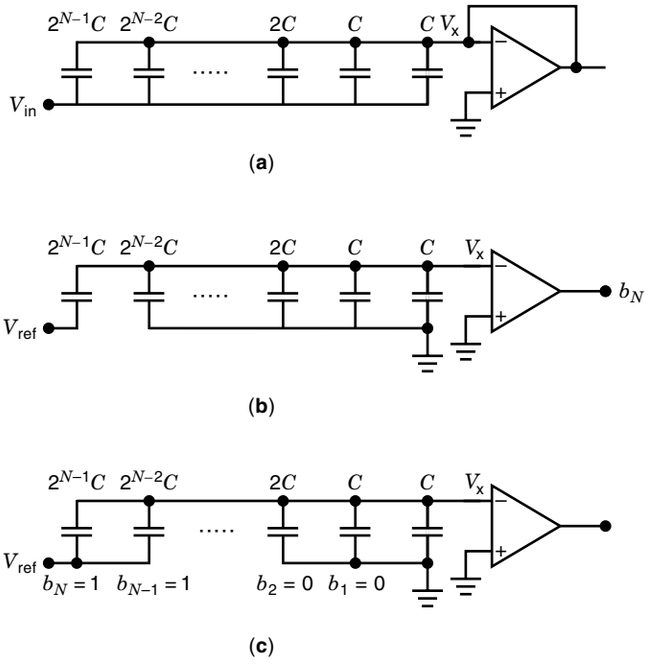


Figure 15. Sequence of configurations of the SC successive approximation ADC during one cycle of the conversion algorithm.

In the latter case, the MSB is set to '0' and the corresponding branch is connected back to ground. In this case, there will be a further charge redistribution in the array such that the original input voltage appears again at the negative terminal of the opamp. After all bits have been tested by similar sequences of charge redistribution and comparison, the connection of the capacitors either to ground or to the reference voltage will reflect the final configuration of the converted output digital word, as seen in Fig. 15(c).

In the preceding SC successive approximation ADC the accuracy of conversion is limited both by capacitance mismatch errors and by the input referred offset voltage of the opamp. Capacitance matching accuracy can be maximized by adopting careful layout designs of the capacitor arrays properly sized, whereas the input-referred offset of the opamp can be virtually eliminated by means of auto-zero techniques. Thus, untrimmed conversion resolutions of up to 9 to 10 bits can be comfortably achieved by SC successive approximation ADCs. For higher resolutions it is mandatory to employ calibration techniques which can extend the inherent matching accuracy of integrated capacitor-arrays well above the inherent technology-dependent 10-bit level.

Regarding the speed of conversion, there are two main limitations imposed on the operation of the SC successive approximation ADC in Fig. 14. One is due to the speed of the comparator, whereas the other results from the *RC* time constants associated with each SC branch because of the equivalent on-resistance of the switches. In most cases such time constants can be properly controlled by sizing accordingly both the transistor aspect ratios and the input capacitors and hence the speed of conversion becomes limited essentially by the speed of the comparator. This allows the design of SC successive approximation ADCs in modern CMOS technology

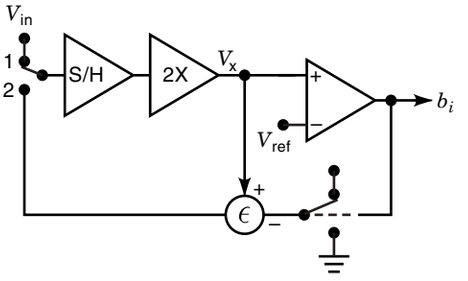


Figure 16. Conceptual block diagram of a multiply-by-two cyclic ADC.

to meet a wide range of conversion specifications, from a few kilohertz to the megahertz range.

Cyclic ADC. Another type of algorithmic ADC is illustrated in the conceptual diagram shown in Fig. 16, comprising an input sample-and-hold, a two-fold gain amplifier, a comparator, and an adder (7). In the sampling phase, the input multiplexer is initially connected to the input terminal. Next, the input signal is multiplied by two and compared with V_{ref} to resolve the most significant bit. Depending on the result, a residual signal will be fedback to the input of the circuit for further processing (thus the designation of cyclic conversion). If $V_{in} > V_{ref}$, then MSB = 1 and the generated residual signal will be $(2V_{in} - V_{ref})$. If, on the contrary, $V_{in} < V_{ref}$, then MSB = 0 and the residual signal will be simply the signal $2V_{in}$. The residual signal will be again multiplied by two and compared with V_{ref} to resolve the second most significant bit and again generate a new residue. This cycle repeats as many times as the number of bits to be resolved and, hence, the conversion time is directly proportional to the resolution of the converter.

A possible SC implementation of a cyclic ADC is shown in Fig. 17. The first opamp samples the input signal in the first conversion cycle, and the residue signal in the remaining conversion cycles. The second opamp performs the combined functions of amplification by two of the recycling residue and addition. The 1-bit digital-to-analog conversion function is performed by the switched capacitor that can be connected either to V_{ref} or ground. The number of components and capacitance spread is independent of the conversion resolution and,

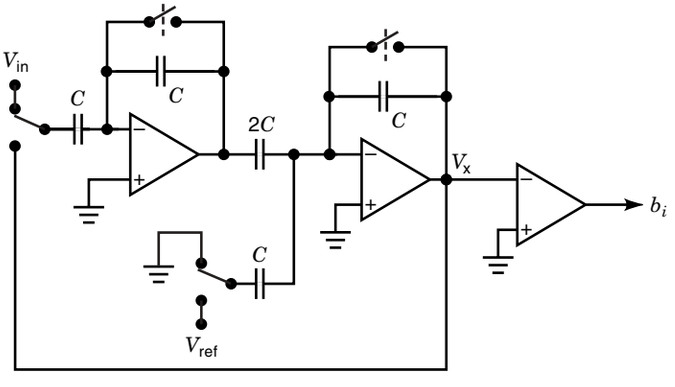


Figure 17. Switched-capacitor realization of a cyclic ADC.

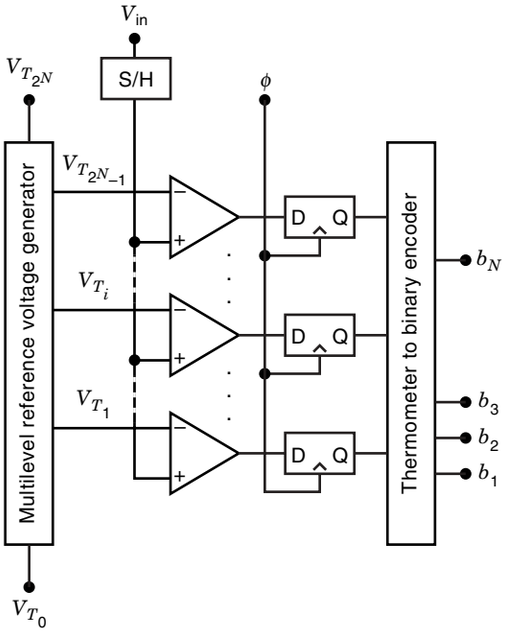


Figure 18. Typical block diagram of a flash ADC.

hence, the converter can be very compact even for medium to high resolution.

Parallel Processing ADCs (8)

Flash ADC. The block diagram of a flash ADC is represented in Fig. 18 (9). It comprises an input sample-and-hold circuit, a multilevel reference voltage generator with output taps directly connected to the inputs of a bank of comparators and, finally, some digital circuits for latching and encoding the output voltages of the comparators. The input voltage is sampled and held and then compared simultaneously with $2^N - 1$ voltages (for a resolution of N-bit) of the reference voltage generator. The resulting thermometer code at the output of the comparators is encoded into an N-bit two's-complement binary code and then latched.

One circuit solution that is particularly suitable for integrated circuit implementation in CMOS technology is shown in Fig. 19. Here, the multilevel reference voltage generator is formed by a resistor-string with direct connection of the output voltage taps to the inputs of a bank of latched comparators. Each of the latched comparators in the converter employs input offset voltage compensation by auto-zero. Slight modifications of this architecture can also be considered if a specialized sample-and-hold circuit is employed, rather than the distributed scheme already mentioned here.

With only one clock cycle per conversion period, the flash ADC gives the fastest possible time of conversion. However, because of the very large number of resistances and especially comparators needed (proportional to 2^N), this type of converters usually occupies a very large area for implementation and burns a significant amount of power. In practice, they are therefore limited to conversion resolutions of no more than 8-bits. In order to overcome such limitations and still achieve high speed of conversion, two alternative types of parallel processing ADCs are described next, namely the subranging and the two-step flash converters.

Subranging ADC. The conceptual block diagram of the subranging ADC is illustrated in Fig. 20 and which comprises, besides the input sample-and-hold, two flash-like converters, one with N_1-bits and the other with N_2-bits, connected by a digital-to-analog conversion block (10). The resulting digital codes of both flash converters are combined to form the converted output digital word of the converter. After input signal sampling, the operation of the subranging converter is performed in two phases, namely a coarse conversion phase where typically the N_1 most significant bits are determined, and a fine conversion phase where the remaining N_2 least significant bits are obtained.

The behavior of the subranging converter in both the coarse and fine conversion phases is identical to the behavior of the flash, although the set of voltages to which the input voltage is compared is different from one phase to the other,

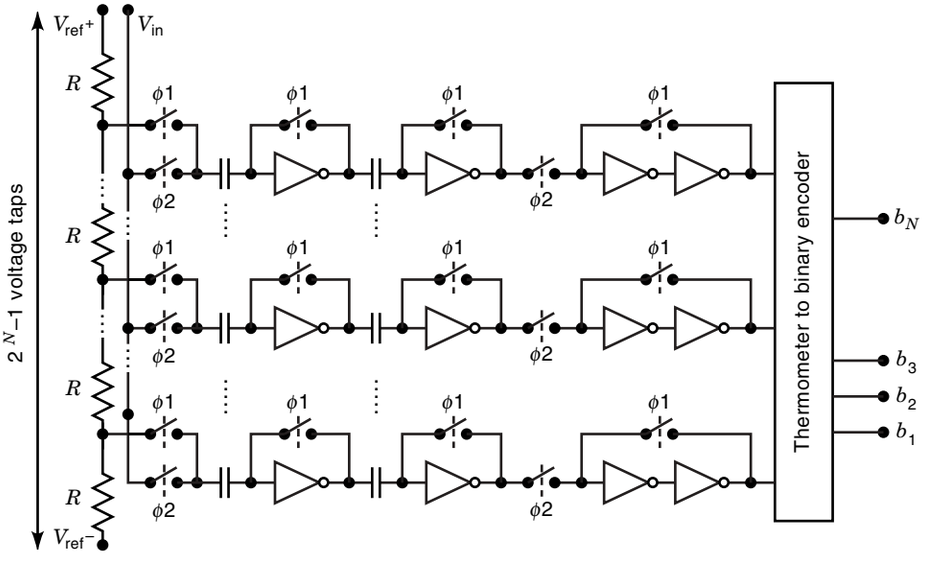


Figure 19. The CMOS implementation of a flash ADC.

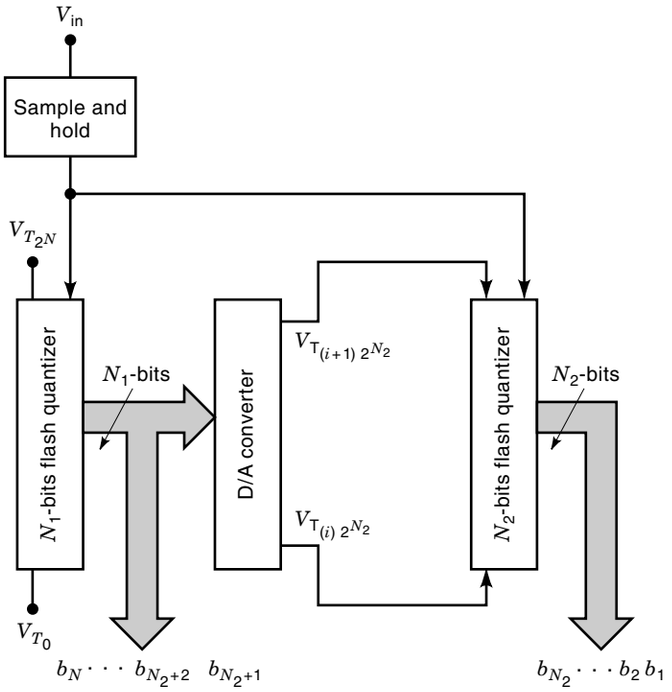


Figure 20. Conceptual block diagram of a subranging ADC.

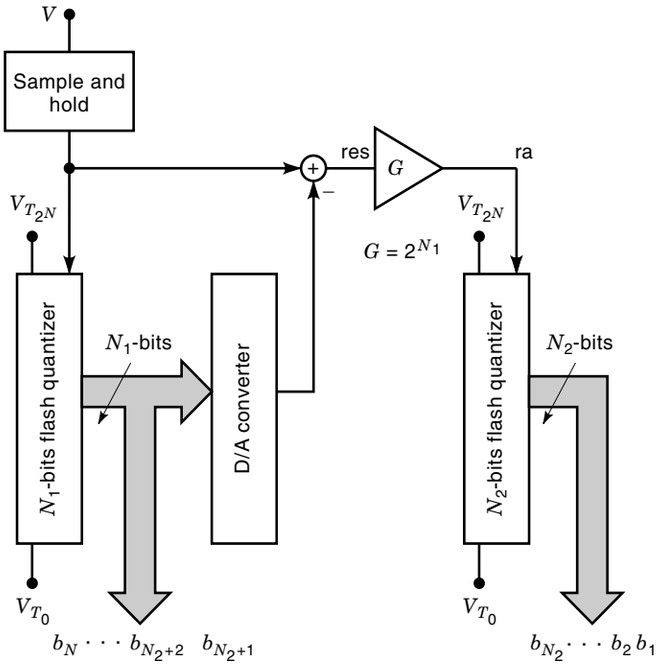


Figure 22. Conceptual block diagram of a two-step flash ADC.

as seen in Fig. 21. In the coarse conversion phase the input voltage is compared with a set of $2^{N/2} - 1$ voltages spanning the whole input signal range, and the resulting MSBs indicate the specific conversion interval that encompasses the sampled input voltage. Then, in the fine conversion phase the selected conversion interval is segmented into a further $2^{N/2} - 1$ voltages in order to determine the remaining LSBs.

In subranging ADCs the required number of comparators and resistors is only proportional to $2^{N/2}$ rather than to 2^N as in the case of the pure flash ADC. Therefore subranging ADCs offer an attractive solution for the integrated implementation of high-speed ADCs with 8-bits resolution and above.

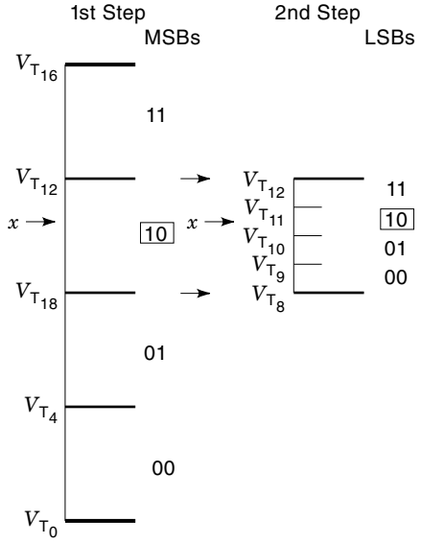


Figure 21. Illustrating the two quantization phases in a subranging ADC.

Two-Step Flash ADC. The conceptual block diagram of a two-step flash ADC is shown in Fig. 22, comprising two flash quantizers, one with N_1-bits and the other with N_2-bits, a digital-to-analog converter (DAC) and also a gain block that amplifies the difference between the input analog signal and the partially reconstructed analog signal obtained at the output of the DAC (11,12). The conversion cycle is divided into a sampling phase, a coarse conversion phase, and a phase for the amplification of the residue and its fine conversion. After sampling, the input voltage is applied to the input of the N_1-bits coarse quantizer to determine the MSBs that are latched into a digital register. Afterwards, the converted digital code is reconstructed back to an analog voltage by means of the DAC. This voltage is then subtracted from the input voltage signal to form the residue signal of the converter. This, in turn, is amplified by the gain block $G = 2^{N_1}$ and then applied to the N_2-bits fine quantizer to determine the LSBs of the converter. The final output digital code is obtained by combining the MSBs of the first quantization phase together with the LSBs of the second quantization. The accuracy of the amplified residue signal generated internally in the two-step flash ADC is a key factor to determine the overall linearity of the converter. Under ideal circuit conditions it evolves between two consecutive digital codes as illustrated in Fig. 23, where the slope of the characteristic is defined by the amplification factor of the gain block.

A circuit implementation of the two-step flash ADC that is particularly suitable for integration in CMOS technology is shown in Fig. 24. The multiplying DAC uses an array of switched capacitors and an opamp, whereas the flash quan-

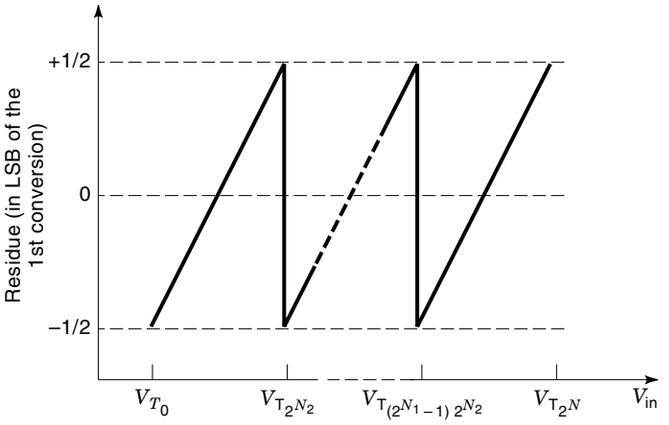


Figure 23. Amplified residue signal internally generated in the two-step flash converter under ideal circuit components characteristics.

tizer is identical to the one presented in Fig. 19 because it first samples the voltages generated by the resistor-string into the input capacitors (C_i), and then compares them with the input voltage. It is possible to realize both the coarse and fine conversions by the same flash. The operation of this particular configuration is as follows. The input voltage is sampled in the top plates of the capacitor array, while the input capacitors of the flash sample the corresponding voltages generated in the resistor-string. In the next phase, the whole array is connected as the feedback of the opamp, thus holding at its output the sampled voltage. This output is next applied to the input of the flash to determine the $N/2$ MSBs. In the next phase, the resolved bits are applied to the DAC to subtract the reconstructed analog voltage from the original sampled voltage. A residue amplified by $2^{N/2}$ is then generated at the output of the amplifier. This is again applied to the flash to determine the remaining $N/2$ LSBs.

Digital Error Correction. Because of unavoidable circuit impairments associated with practical integrated circuit realizations, the quantization window determined by the first flash quantizer in both the subranging and two-step flash ADCs is determined with a maximum error of $\pm 1/2$ LSB of its nominal resolution, rather than with an ideal 0 LSB error. When this error is passed on to the second quantization phase it may overflow the quantization range of the flash and thus produce errors in the overall conversion characteristic.

To correct for such errors the flash quantizer may be augmented beyond its nominal resolution, for example by $+1/2$ LSB on the upper boundary of the quantization window and by $-1/2$ LSB on the lower boundary of the quantization window, such that the additional resolved bit can be subsequently used for digital code correction. This technique is illustrated in Fig. 25. The quantization zone A corresponds to the original comparators in the center of the flash quantizer, zone B corresponds to the additional lower comparators and zone C corresponds to the additional upper comparators. When the MSB flash quantizer determines correctly the quantization zone A, as in Fig. 25(a), the output digital code is obtained directly from the outputs of the comparators. When the flash quantizer resolves instead a code within zone B, as shown in Fig. 25(b), the output digital code of the previous quantization will be corrected by -1. Finally, when the flash quantizer resolves a code within zone C, as illustrated in Fig. 25(c), the output digital code of the previous quantization will be corrected by $+1$.

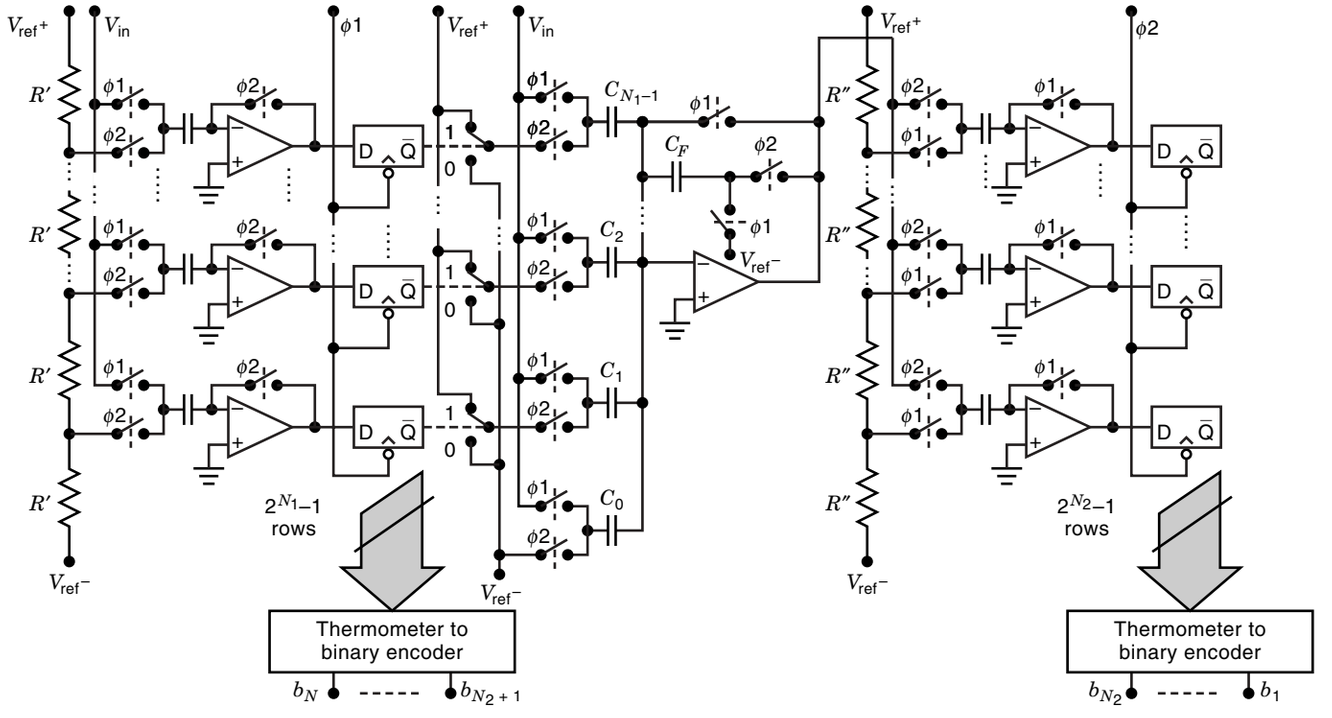


Figure 24. The CMOS implementation of a two-step flash ADC.

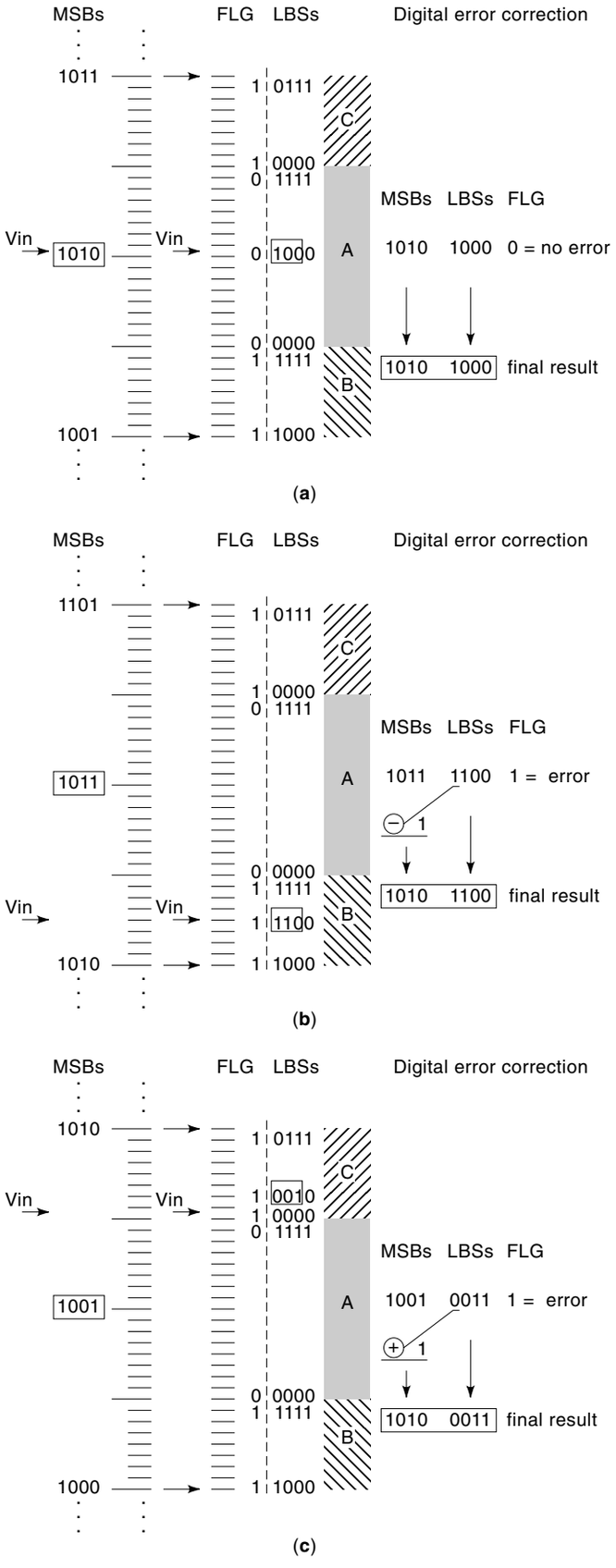


Figure 25. Illustrating the digital error correction technique in an ADC.

Pipeline A/D Converters

Architecture and Operation. Pipelined ADCs constitute a type of parallel processing converters that have gained popularity in a variety of high-speed conversion applications due to the cost/benefit they can achieve over the subranging and two-step flash converters (13,14). The conceptual block diagram of a pipeline ADC is shown in Fig. 26(a), for an example with five pipelined stages, and the associated timing diagram is given in Fig. 26(b). During the first phase of the first clock cycle, the input stage samples the input signal and executes the first quantization. Then, in the second phase the residue is generated, amplified, and transmitted to the second stage of the pipeline for further processing. In subsequent clock cycles similar operations of sampling, quantization, and residue generation are again carried out in the first stage. Meanwhile, in the second stage, these operations are performed in opposite phases. Input sampling and quantization are performed when the first stage is amplifying the residue, whereas residue amplification and transmission to the next stage are performed when the first stage is receiving a new input sample. In parallel with this form of horizontal propagation of the signal (actually, the signal residues), the digital words quantized in each stage and in each clock cycle are propagated vertically through digital registers, as seen in Fig. 26(a), such that at the end of five clock cycles they are all available at the output of the converter to produce the converted digital word. Then, after this latency period, which is needed to fill all the vertical digital registers, there will be one converted digital word every clock cycle.

The maximum resolution achieved with such an architecture is limited mainly by thermal noise, the nonlinearity of the MDACs produced by capacitor mismatches and also by the residue amplification error that is due to the amplifier nonidealities. The error from the flash quantizer can be digitally corrected if kept within $\pm 1/2$ LSB of the nominal resolution of the quantizer and redundancy is used for digital error correction.

Time-Interleaved Converters

An extended concept for parallel processing conversion is based on the use of the type of time-interleaved architectures illustrated in Fig. 27(a) (15,16). These are formed by M paralleled channels whose operation is multiplexed in time, as illustrated in the timing diagram depicted in Fig. 27(b). Each channel in a time-interleaved converter can be formed by any of the types of ADCs previously described. At any given clock cycle, the input and output multiplexers connect only one ADC channel between the input and output terminals, so that one converted digital word is delivered at the output and a new sample of the input signal is taken for conversion. In the next clock cycles the same operation is sequentially repeated with the remaining ADC channels while the first channel carries out the signal conversion; everything should be completed only after the M channels have been served by the multiplexers. Thus, for a given conversion frequency F_s of the ADCs the time-interleaved operation allows to achieve an overall conversion frequency of MF_s.

A variation of the basic time-interleaved ADC architecture based on a quadrature-mirror filter (QMF) bank is indicated in Fig. 28 (17). At the input there is an analysis filter bank, typically realized in switched-capacitor form, whereas the output filter bank is a digital network. In this approach the

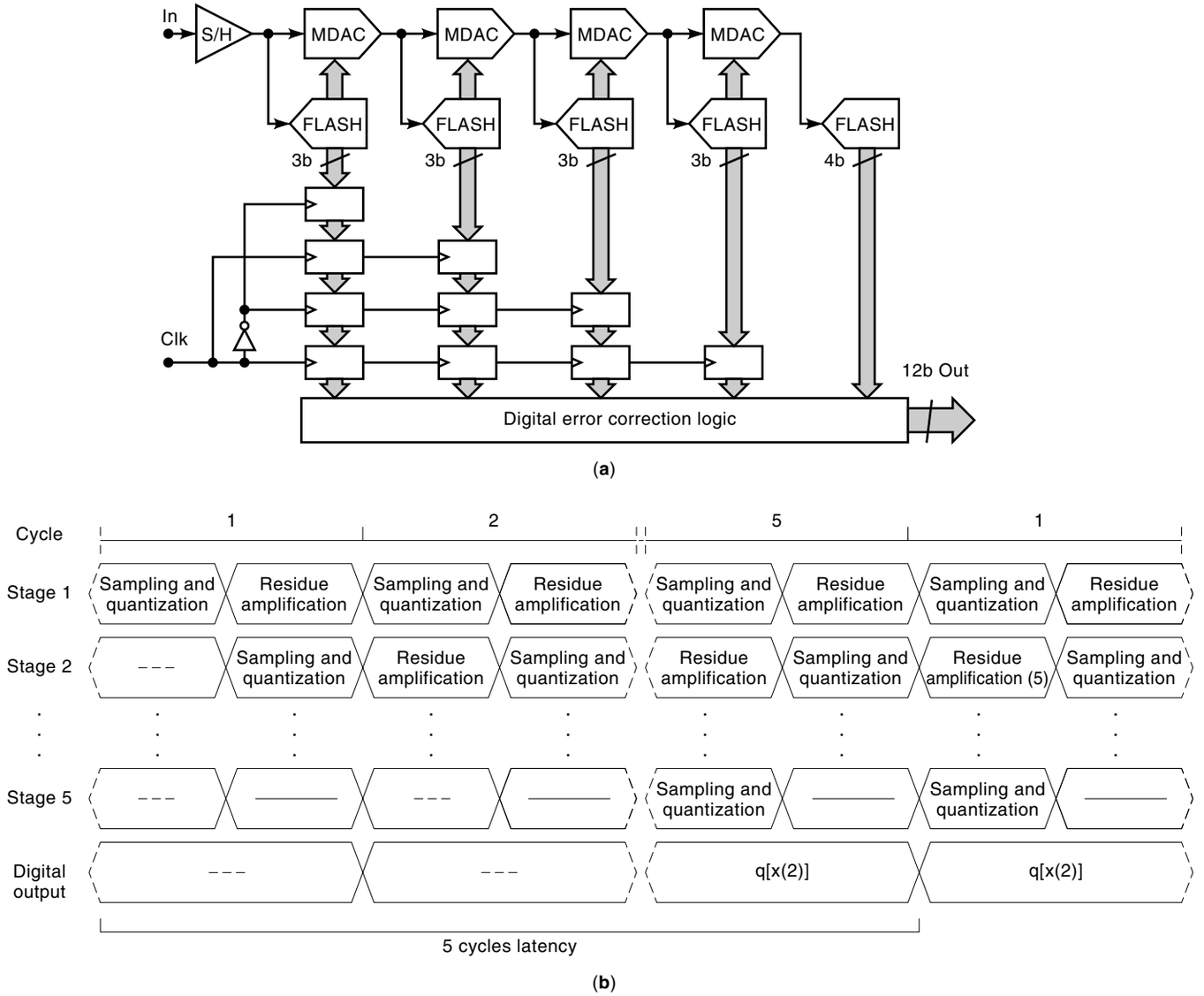


Figure 26. (a) Conceptual block diagram of a five-stage pipelined ADC. (b) Illustration of the timing of the pipe operation and latency of the converter.

input signal is first decomposed into a number of contiguous frequency bands (subbands) so that a specific ADC (subconverter) can be assigned to each subband signal. Due to the synthesis filters the linearity performance due to mismatches among the subconverters is substantially reduced. Similarly, the jitter problem that arises in the basic time-interleaved converter due to uneven sample timing, especially for high-frequency input signals, is also reduced by the filtering and downconversion stage. Besides, this type of QMF-based converter also incorporates the advantages of subband coding such that by appropriately specifying the resolution of the subconverters throughout the respective subbands the quantization noise can be separately controlled in each band, and the shape of the reconstruction error spectrum can be controlled as a function of the frequency.

Theoretically, the overall resolution of a QMF-based ADC depends solely on the resolution of the subconverters used. If successive approximation subconverters are used, then substantial savings in die area can be obtained when compared to flash converters.

OVERSAMPLING CONVERTERS

The common characteristic of all the converters already described herein concerns the uniform power spectral density of the quantization noise in the Nyquist band, that is, the frequency range from dc to half the conversion frequency. Hence, their designation of Nyquist converters. In order to achieve high SNR, Nyquist converters must have a correspondingly high conversion resolution, which, in turn, requires very high matching accuracy from the constituting elements responsible for the scaling operations in the converter. This can be achieved without too much difficulty for conversion resolu-

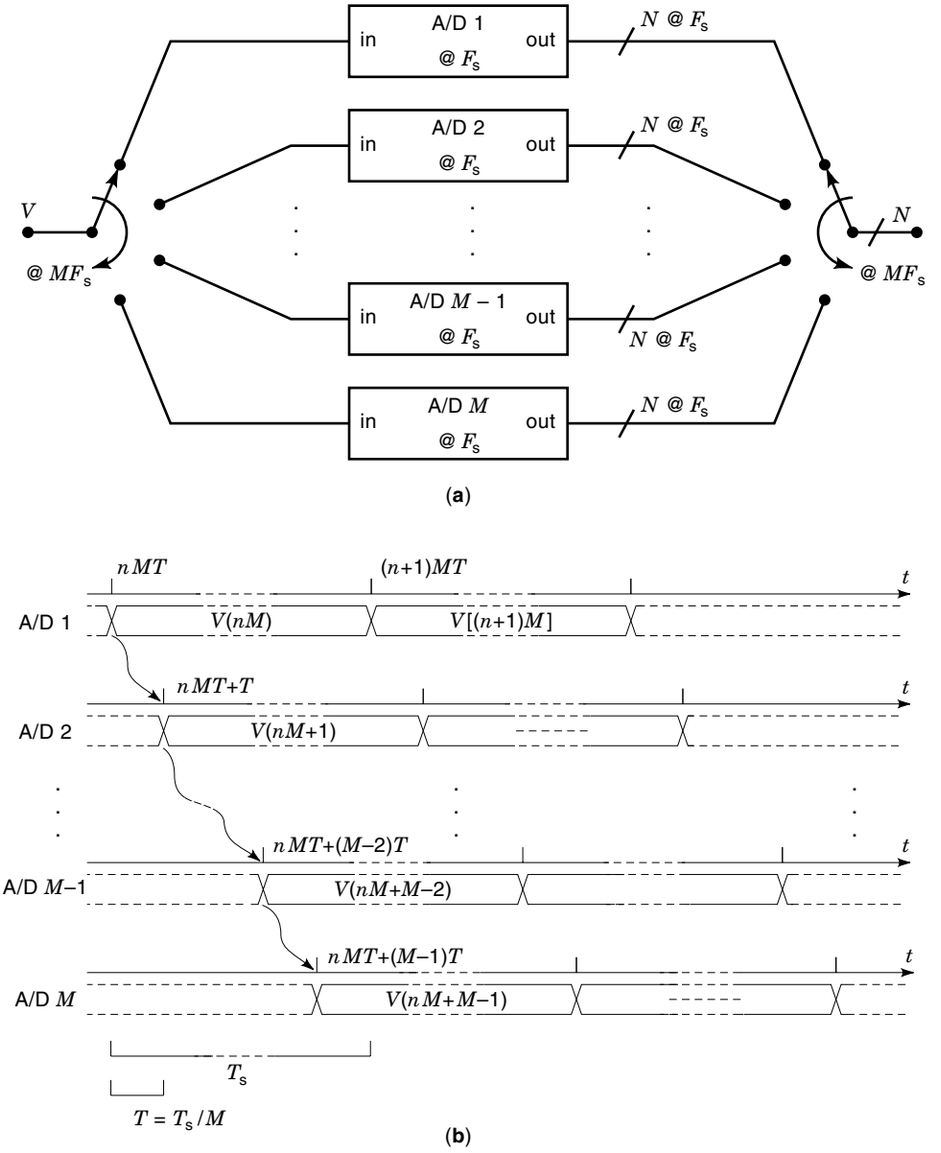


Figure 27. (a) Conceptual block diagram of a time-interleaved ADC and (b) illustration of its time multiplexed operation.

tions up to about 10-bits. For conversion resolutions above 10-bits self-calibration techniques should be employed to extend the matching accuracy of the elements above the inherent accuracy provided by the technology.

Because of the increased difficulty of designing self-calibrated converters alternative techniques have been sought to achieve equivalent high-resolution conversion within the limits of matching accuracy that can be achieved with a given technology. Such conversion techniques are based on shaping the quantization noise of an oversampled signal, that is, where its sampling frequency is much higher than the Nyquist frequency, such that the resulting power spectral density is significantly reduced within the frequency band of interest and increases outside such a band. Hence, the resulting converters are commonly known as oversampling converters.

The notion of using such artificially high sampling rates and simple single-bit quantization to represent analog signals has been of interest ever since the introduction of delta modulation. However, the oversampling technique alone would require sampling frequencies too high to be of much practical use. Subsequent developments introduced the methods of negative feedback, noise shaping and higher-order modulators (18). These improvements allowed the practical implementation of very high resolution converters at the expense of an increase in complexity of the digital filters that are needed to extract the baseband signal from the high speed bit-stream produced by the modulator (19). Because these can be implemented rather efficiently by modern CMOS technology, oversampling converters have been making inroads in mixed-signal analog-digital integrated circuits for high performance applications. The sections that follow describe the basic concepts and most relevant implementation techniques of oversampling ADCs.

Quantization Noise with Oversampling and Shaping

To reduce the in-band noise power of an oversampled quantized signal beyond the value obtained for the uniform noise

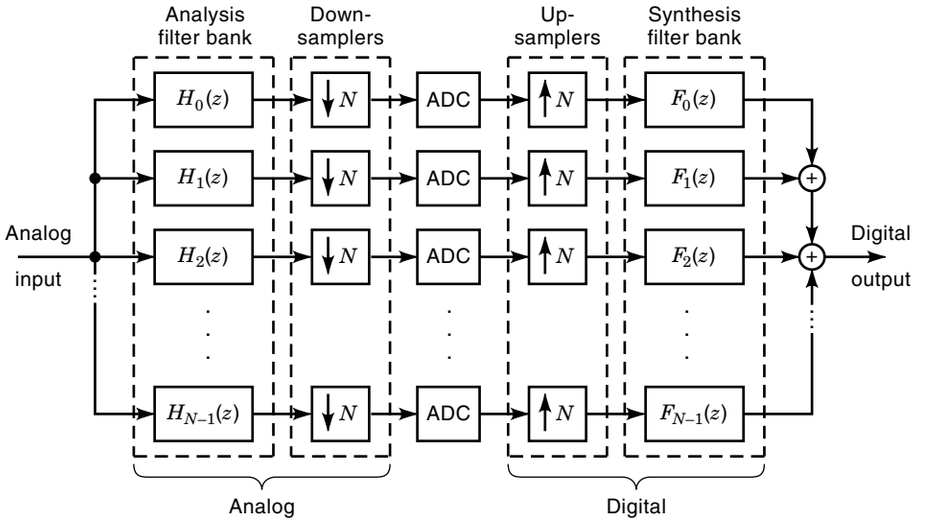


Figure 28. Block diagram of a time-interleaved ADC based on a quadrature-mirror filter (QMF) bank.

distribution, a specific noise shaping function should be performed, as illustrated in Fig. 29. A portion of the quantization noise is now pushed into higher frequencies, thereby improving the equivalent bit-resolution in the signal bandwidth of interest while keeping the oversampling ratio constant. Such shaping effect can be obtained by means of the circuit shown in Fig. 30 for an example in which an integrator is used to produce a first-order noise shaping. Its operation can be intuitively understood from the theory of closed loop systems. Because in the baseband the integrator has a very large gain, the overall transfer characteristic is determined by the feedback branch and any nonlinearity in the feed-forward branch is attenuated by the large loop gain. The feedback branch consists of a digital-to-analog converter and the overall resolution will be determined solely by the linearity of this component.

An interesting case is when only two quantization levels are used, that is, one-bit resolution. This is called a single-bit first-order sigma-delta ($\Sigma\Delta$) modulator and is represented by the circuit diagram shown in Fig. 31. It comprises an integrator for the filter function, a comparator with latch for the 1-bit quantizing function, and a switch to either $-V_{\text{ref}}$ or $+V_{\text{ref}}$ that realizes the 1-bit digital-to-analog conversion providing the reconstruction of signal prior to the subtraction from the incoming signal at the input of the modulator. Because of its simplicity and use of 1-bit converters the system becomes very robust and precise. In particular, the implementation of the 1-bit DAC renders such a structure inherently linear, possibly yielding only offset and gain errors. Thus, the $\Sigma\Delta$ modulator offers the potential for high resolution conversion without the need for accurate components.

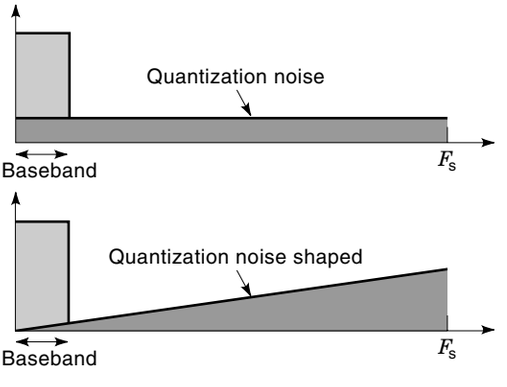


Figure 29. Effect of shaping the quantization noise.

Analysis of First- and Second-Order $\Sigma\Delta$ Modulators

Due to the highly nonlinear element in the loop, the exact analysis of a $\Sigma\Delta$ modulator is not simple to do and any form of analytical solution is too complex to be of much practical use. Therefore, a simplified linearized model of the modulator is used in which the comparator is replaced by an additive noise source. Although this approximation allows prediction of some important aspects of the modulator behavior, such as noise level and spectrum, it must be nevertheless carefully interpreted because it assumes a set of conditions not thoroughly satisfied in most applications.

First-Order $\Sigma\Delta$ Modulator

Linearized Model. Referring to the circuit in Fig. 31, the corresponding linearized model of the modulator is represented in Fig. 32, where the integrator is modeled by a discrete-time function and the quantizer is modeled as a unity gain block with an additive noise source V_{Q}. It can be readily obtained that the transfer function for the input signal is given by

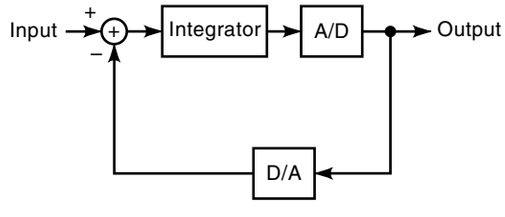


Figure 30. Oversampling and noise-shaping modulator.

$$S(z) = \frac{V_{\text{out}}(z)}{V_{\text{in}}(z)} = z^{-1} \tag{21}$$

whereas the transfer function relative to the quantization noise source is given by

$$Q(z) = \frac{V_{\text{out}}(z)}{V_{\text{Q}}(z)} = (1 - z^{-1}) \tag{22}$$

to yield the frequency response

$$Q(f) = 2 \sin\left(\frac{\pi f}{F_{\text{s}}}\right) \tag{23}$$

Thus, while the input is unaffected throughout the modulator processing chain (apart from the delay term), the quantization noise V_{Q} is filtered by a high-pass function with notch frequency at dc.

Signal-to-Noise Ratio. For the circuit considered above with the noise related transfer function given by Eq. (23), the total quantization noise power in a frequency band from dc to f_{c} can be calculated from

$$P_{\text{Q}} = \int_0^{f_{\text{c}}} \left(\frac{\text{LSB}^2}{12}\right)\left(\frac{1}{F_{\text{s}}/2}\right)\left[2 \sin\left(\frac{\pi f}{F_{\text{s}}}\right)\right]^2 df \tag{24}$$

where, as previously indicated, LSB represents the quantizer step size.

Considering that the signal band of interest is highly oversampled, that is, $(f_{\text{c}}/F_{\text{s}}) \ll 1$, the total in-band noise power can be expressed as follows:

$$P_{\text{Q}} = \left(\frac{\text{LSB}^2}{12}\right)\left(\frac{\pi^2}{3}\right)\left(\frac{1}{\text{OSR}}\right)^3 \tag{25}$$

as a function of the quantization step and oversampling ratio of the modulator. Hence, for a sine waveform with maximum peak value of $(2^N - 1)(\text{LSB}/2)$ it results in a signal-to-noise ratio given by

$$\text{SNR} = 10 \log\left[\frac{3}{2}(2^N - 1)^2\right] + 10 \log\left(\frac{3}{\pi^2}\text{OSR}^3\right) \tag{26}$$

and which, for a 1-bit quantizer ($N = 1$), can also be expressed as

$$\text{SNR(dB)} = 9 \log_2(\text{OSR}) - 3.41 \tag{27}$$

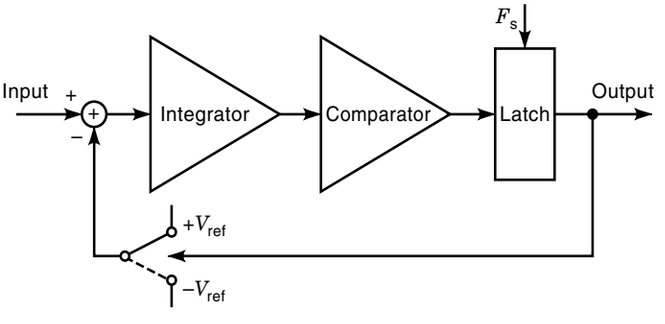


Figure 31. Schematic diagram of the first-order ΣΔ modulator.

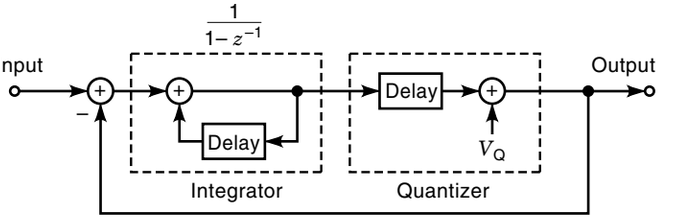


Figure 32. Linearized model of the first-order ΣΔ modulator.

From the preceding expression it is now clear that each octave of the oversampling ratio leads to a 9 dB improvement of the SNR, which corresponds to an equivalent increase of 1.5 bit of the conversion resolution.

One important aspect in the type of oversampling modulator considered above, and that is not predicted by the approximated linearized model, is the appearance of limit cycles for certain input waveforms (20). The reason is that the quantization noise becomes highly correlated with the input. Then, the assumption that the quantization noise spectrum is white is no longer valid. Actually, the quantization noise is concentrated in discrete frequency bands that, when falling in the baseband, produce noise tones much above the noise level predicted by Eq. (25). This effect can be attenuated by introducing a high frequency dither signal superimposed on the input signal, thereby creating sufficient disturbance so as to destroy the tones. However, the addition of this dither signal reduces the dynamic range of the modulator and complicates the design.

Second-Order ΣΔ Modulator. By adding another integrator in the forward loop, a stronger reduction of low-frequency quantization noise is possible for the same oversampling ratio. Furthermore, due to the additional integrator the quantization noise becomes a more complex function of the circuit parameters and it is, therefore, less correlated with the input. Thus, a second-order modulator will be much less prone to enter in limit cycle conditions than its first-order counterpart and the corresponding noise tone power is very small. In applications where the input signal is sufficiently busy so as to completely randomize the quantization noise there is no need to add a dither signal. Figure 33 shows the schematic diagram of such a second-order oversampling modulator. The inner second feedback loop has been added to ensure operation stability.

Two of the most common implementations of second-order modulators are represented by their linearized discrete-time models illustrated in Fig. 34. The first one, represented in Fig. 34(a), employs two delay-free integrators whereas the other one, represented in Fig. 34(b), employs a delayed integrator in the first stage and a different coefficient in the inner loop. The latter implementation allows more design flexibility due to the relaxed timing requirements, and smaller voltage swings on the integrator outputs (21).

From the linearized models represented in the preceding illustration (Fig. 34), it is readily seen that the signal is merely affected by a delay term corresponding to one clock period, in the case of the delay-free integrators, and two clock periods in the other case. Both forms of implementation produce the same noise shaping effect determined by a high-pass

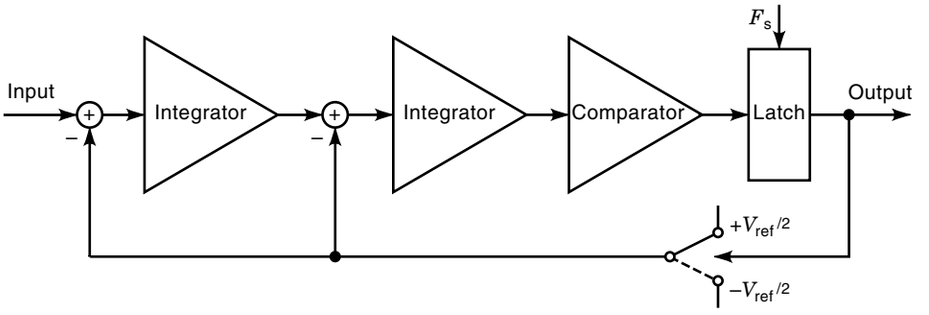


Figure 33. Schematic diagram of a second-order ΣΔ modulator with two feedback loops for operation stability.

transfer function expressed as

$$Q(z) = (1 - z^{-1})^2 \tag{28}$$

yielding the frequency response

$$Q(f) = \left[2 \sin\left(\frac{\pi f}{F_s}\right)\right]^2 \tag{29}$$

It can be appreciated that while at low frequencies the quantization noise will be strongly attenuated due to the second-order noise shaping the high-frequency portion of the spectrum will be substantially amplified, as is illustrated in Fig. 35.

Following a similar procedure as for the first-order modulator, it can be shown that in both forms of implementation the resulting quantization noise power at the output, in a frequency band from dc to $f_c = F_s/2\text{OSR}$, is given by the approximate expression

$$Q \cong \frac{\text{LSB}^2 \pi^4}{60} \left(\frac{1}{\text{OSR}}\right)^5 \tag{30}$$

and which for a single-bit quantizer yields

$$\text{SNR(dB)} = 15 \log_2(\text{OSR}) - 11.14 \tag{31}$$

Therefore, each doubling of the oversampling ratio provides a 15 dB increase in the SNR, which gives an equivalent resolution increase of 2.5 bits.

Figure 36 compares the SNR characteristics obtained for oversampling ΣΔ modulators with first- and second-order filtering as well as for oversampling modulators with no noise shaping function (zero-order). For example, in order to achieve an equivalent 12 bit of resolution with a second-order oversampling modulator a signal-to-noise ratio of 74 dB must be obtained. In practical application designs, in order to allow for secondary sources of noise such as thermal noise in the input stage and degradation of noise shaping due to imperfect components, a margin of approximately 6 dB should be added. From Fig. 36 we can see that for obtaining SNR = 80 dB such a second-order modulator would need an OSR of only 65, whereas a first-order modulator would required a much higher OSR of 600.

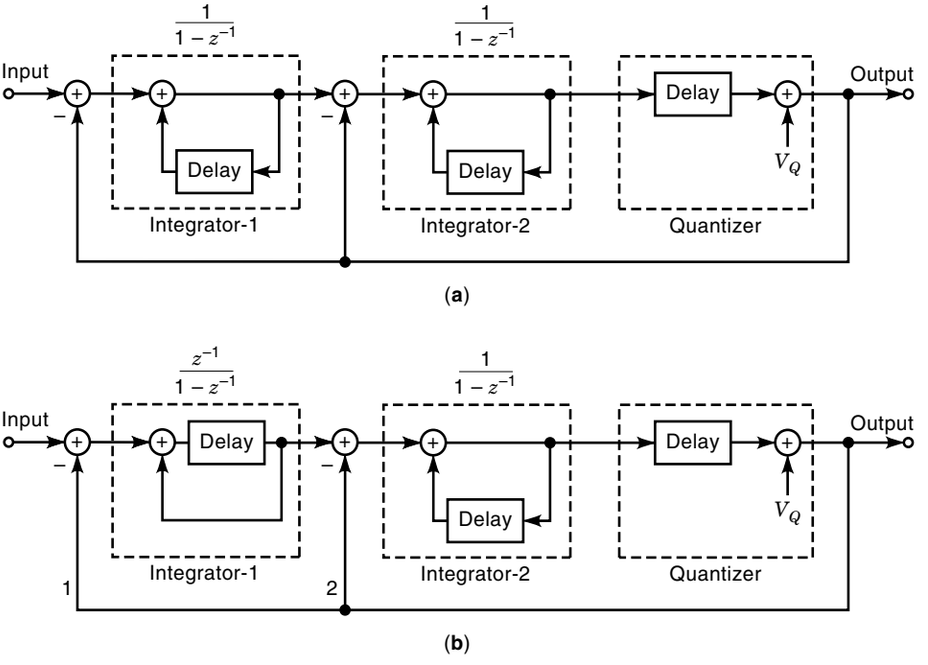


Figure 34. Linearized equivalent circuits of second-order ΣΔ modulators for two of the most popular implementations.

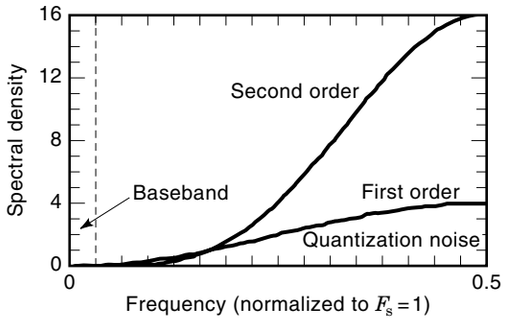


Figure 35. Shaping functions of the output quantization noise for the case of first- and second-order ΣΔ modulators.

System Aspects of Oversampling ADCs

The general architecture of a complete ADC based on oversampling techniques is represented in Fig. 37. Besides the analog modulator clocked at a high (oversampled) frequency, the system includes a digital filter, also clocked at the oversampling frequency, and an output register clocked at the lower Nyquist frequency. The combined digital filter and output register with different clock rates perform the so-called decimation function of the converter, which purpose is to remove the high-frequency noise components produced by the shaping effect of the modulator.

Figure 38 depicts several time-domain waveforms and corresponding signal spectra that illustrate the operation of the complete oversampling ADC. It is assumed that an input signal band limited from dc to $f_c \ll F_s$ is sampled at the high-frequency clock, as seen in Fig. 38(a). Next, the 1-bit quantized bit stream at the output of the modulator contains basically the baseband information from dc to f_c and a large amount of out-of-band quantization noise above f_c. The corresponding time-domain waveform and frequency spectrum are depicted in Fig. 38(b). The out-of-band quantization noise is then removed by means of the digital filter and which, at the same time, increases the length of the digital signal from 1 bit to the full N-bits resolution of the converter. The resulting spectrum is shown in Fig. 38(c). Finally, the output sampling rate is reduced by means of an M-fold down sampler in order to obtain to the required output conversion frequency and thus yields the periodic (digital) spectrum illustrated in Fig. 38(d). This last stage of the processing chain is called decimation.

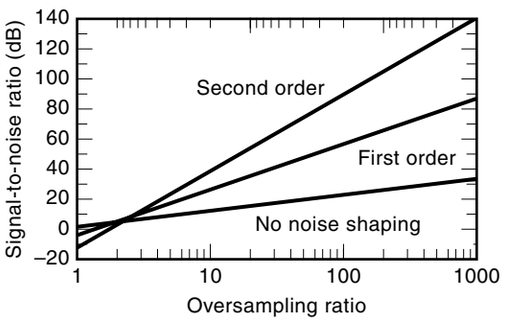


Figure 36. The SNR characteristics of oversampling modulators with zero-order (no noise shaping), first-order, and second-order noise shaping as functions of the oversampling ratio.

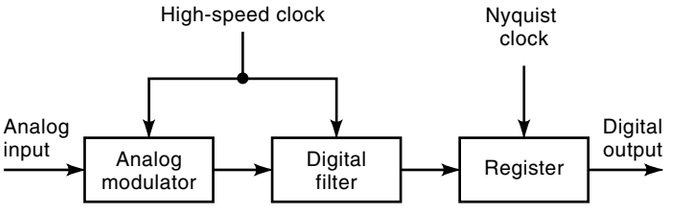


Figure 37. Block diagram of a complete oversampled ADC system, including an analog modulator and a digital decimator.

The design of a complete oversampling ADC based on the preceding system architecture involves the selection of key design parameters for both the analog modulator and the digital decimator, bearing in mind the interaction between them. For the analog modulator, on the one hand, the parameters of concern are the order of the filtering function and the number of quantized bits. On the other hand, the relevant design parameters for the digital decimator are the oversampling ratio, the word length at both the output of the analog modulator and at the output of the system, and the required level of attenuation of the out-of-band noise. Next, we shall discuss more advanced architecture options for designing oversampled ADCs.

Higher-Order Modulators. The oversampling ratio required to meet a specific level of performance may be decreased below that needed in a first-order ΣΔ modulator by increasing the order of the modulator. Higher-order noise shaping can be accomplished by including a higher order filter, such as additional integrators, in the forward path of the modulator. However, higher-order modulators require careful attention to the placement of appropriate zeros in the transfer function of the analog filter (22,23). Moreover, when a higher-order modulator is driven by a large input, the two level quantizer is overloaded, causing an increase in the quantization noise. The increased quantization noise is amplified by the analog filter, leading to instability in the form of large, uncontrollable, low-frequency oscillations. Thus, third- and higher-order modulators based on the use of a single two-level quantizer are potentially unstable and may require circuitry to reset the integrators when large signals are detected in the integrator outputs (24).

Cascaded Modulators. The danger of having saturating limit cycles in high-order modulators can be avoided by cascading a number of first- and second-order ΣΔ modulators to produce the effect of high-order prediction. One such architecture is represented in Fig. 39 showing a cascade of two first-order ΣΔ modulators. In this arrangement the second modulator takes at the input the quantization error of the first stage while the outputs of both modulators are combined together. In the combined output signal the first-stage quantization error is removed, thus leaving only the error corresponding to the second modulator stage. The technique can generally be extended to more than two stages and to both first- and second-order modulators.

The forgoing type of cascade architectures, called MultistAge noise SHaping (MASH) gives the advantage of achiev-

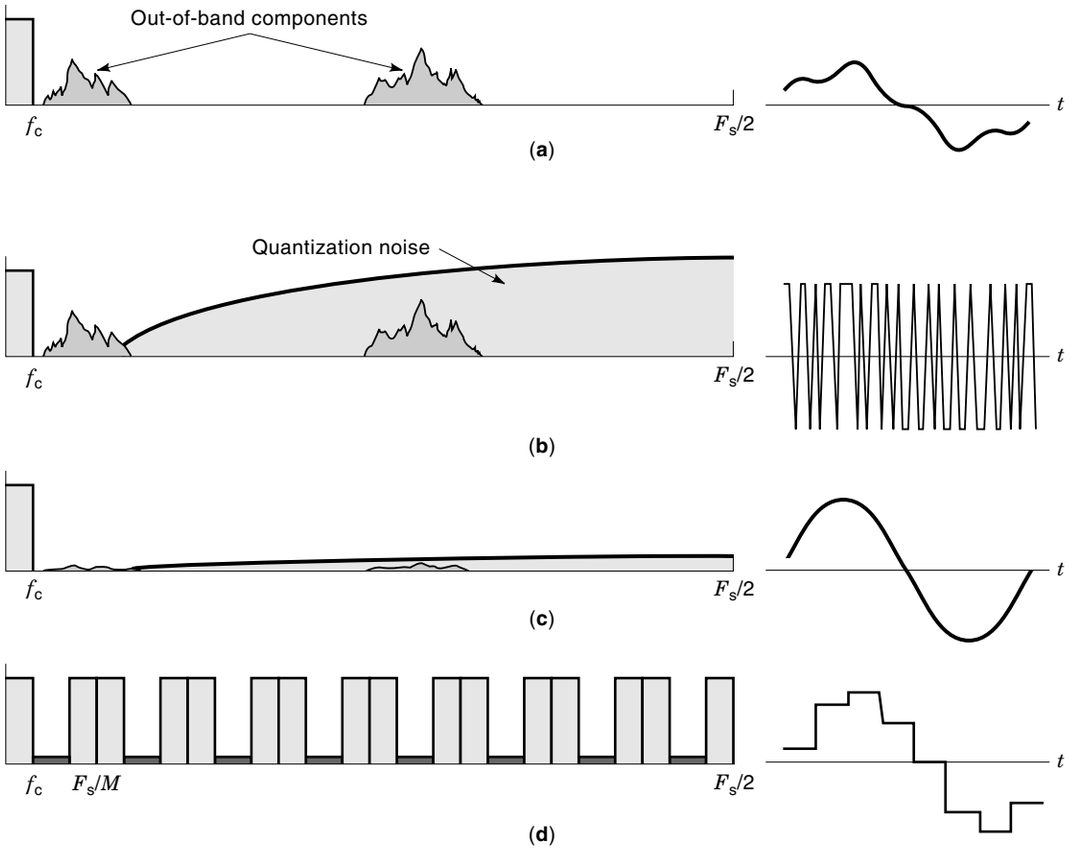


Figure 38. Time-domain waveforms and signal spectra throughout the processing chain of an oversampling ADC. (a) Sampled input analog signal. (b) Digital bitstream signal. (c) Digitally filtered signal. (d) Decimated digital signal.

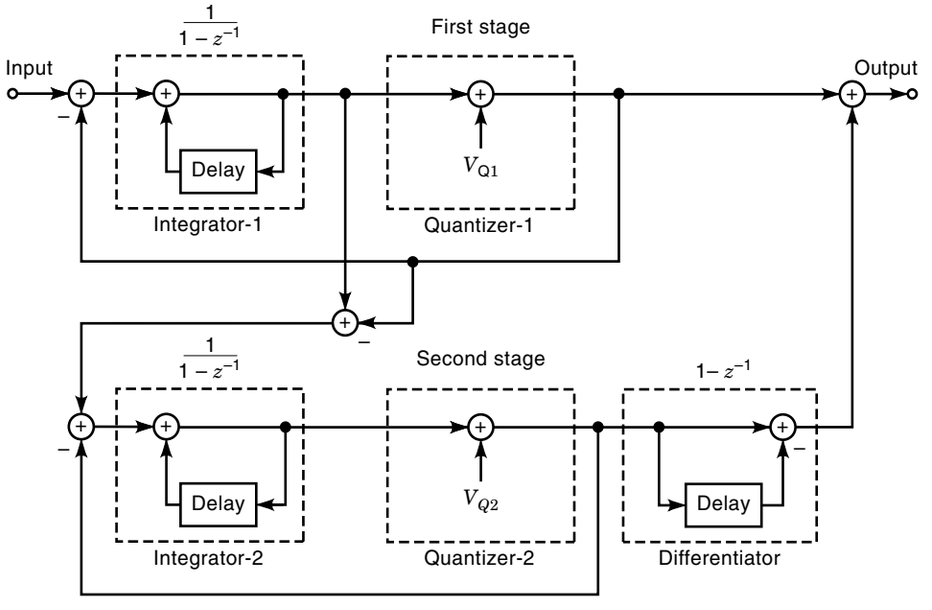


Figure 39. Second-order noise shaping by cascading two first-order modulators.

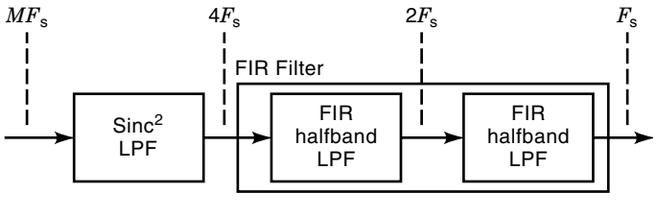


Figure 40. A two-stage architecture for designing the digital decimator using only FIR filter sections.

ing high-order noise shaping functions using inherently stable modulator stages (25–28). Rather than regarding the MASH architecture as a method of obtaining high-order circuits, it is usually more correct to regard it as a means of enhancing the performance of the first modulator in the cascade. For example, a modulator composed of a second-order sigma-delta circuit followed by a first-order circuit has attractive properties. Unfortunately, however, these modulators are sensitive to the opamp finite dc gain as well as to mismatches among the circuit parameter values. Their resolution is usually determined by how well individual modulators are matched. Specifically, in switched-capacitor implementations, cascaded modulators require close capacitance matching, high opamp dc gain, and nearly complete settling of the integrator outputs.

Despite their attractive features, which allow lower oversampling ratios and, therefore, higher conversion rate for a given modulator operating speed, sophisticated modulator architectures such as the ones discussed above do not necessarily ease the performance required on the overall circuit. For example, a larger percentage of the thermal noise introduced by the sampling switches in switched capacitor integrators falls in the baseband. To maintain the dynamic range, the capacitors must be increased accordingly. This implies proportionally higher-load capacitances on the operational amplifiers in the integrators. Also, the complexity of the antialiasing filter that precedes the modulator and the decimator filter following it are increased. Their attenuation specifications are tighter because the sampling rate is lower with respect to the baseband.

***N*-Bit Quantizer.** As previously discussed, the advantage of single-bit modulators is that the linearity of the feedback DAC is inherently ideal, besides being extremely simple for implementation. Single-bit modulators, however, also have the disadvantage of a large amount of quantization noise, which may easily cause saturation and lead to potential instability conditions. Multibit quantizers, by contrast, generally provide improved stability conditions of the $\Sigma\Delta$ modulators, especially in the case of higher order modulators, as well as minimization of the occurrence of idle tones. However, modulators based on a quantizer with more than two levels place stringent linearity demands on the DAC in the feedback loop and generally require sophisticated linearization techniques (29–32).

Decimation Filter. The output of the modulator represents the input signal together with its spurious out-of-band components, quantization noise, and noise or interference that may have entered the analog circuits. As already discussed, the digital filter in the general architecture of Fig. 36 serves to attenuate all out-of-band energy, so that the signal may be resampled at the Nyquist rate without being corrupted with the folded-back components of the high frequency noise.

Fairly simple digital filters would suffice to remove only quantization noise because it rises slowly, for example at 12 dB per octave for the case of a second-order modulation. By contrast, highly selective filters are usually needed to remove the out-of-band components of the input. Such filters are rather expensive when operated at high sample rates. Therefore, in practice, the decimator filter is usually implemented in two stages, as seen in Fig. 40. First, there is a decimator with output at four times the Nyquist rate and which is designed predominantly to remove the quantization noise component that is dominant at high frequencies. The second-stage filter resamples the signal at the Nyquist rate and defines the baseband cut-off characteristics. Typically this is the most complex and larger circuit and should be carefully designed to suit the application.

A convenient filter for the first stage of decimation is based on the Sinc function expressed as

$$\mathrm{Sinc}_k = \left(\frac{1}{M}\frac{1-z^{-M}}{1-z^{-1}}\right)^k \tag{32}$$

where M is the decimation factor and k is the order of the filter. This filter is very easy to implement as it requires no multipliers. It has the advantage of having zeros at the multiples of the output sample frequency, which remove the noise components that would otherwise be aliased into the baseband with the decimation operation (33–35). The order of this filter should be equal to the modulator order plus one in order to suppress the high frequency quantization noise adequately. Eventually, the order of the decimating filter can be made equal to that of the modulator, in order to reduce the implementation complexity. However, this results in some degradation of the overall signal-to-noise ratio. A typical architecture for the implementation of a third-order Sinc filter is illustrated in Fig. 41.

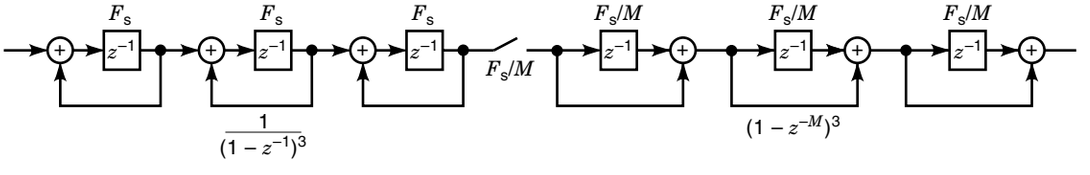


Figure 41. Typical architecture of a third-order Sinc filter for digital decimation.

BIBLIOGRAPHY

1. P. R. Gray and R. R. Neff, Analog-digital conversion techniques for telecommunications applications. In J. E. France and Y. Tsividis (eds.), *Design of Analog-Digital VLSI Circuits for Telecommunications and Signal Processing,* Engelwood Cliffs: Prentice-Hall, 1994.
2. A. Gersho, Principles of quantization, *IEEE Trans. Circuits Systems,* **CAS-25**: 427–436, 1978.
3. S. K. Tewksbury et al., Terminology related to the performance of S/H, A/D, and D/A circuits, *IEEE Trans. Circuits Systems,* **CAS-25**: 419–426, 1978.
4. M. Shinagawa, Y. Akazawa, and T. Wakimoto, Jitter analysis of high-speed sampling systems, *IEEE J. Solid-State Circuits,* **25**: 220–224, 1990.
5. J. Robert et al., A 16-bit low-voltage CMOS A/D converter, *IEEE J. Solid-State Circuits,* **SC-22**: 157–163, 1987.
6. J. Robert and P. Deval, A second-order high-resolution incremental A/D converter with offset and charge injection compensation, *IEEE J. Solid-State Circuits,* **23**: 736–741, 1988.
7. P. Li et al., A ratio-independent algorithmic analog-to-digital conversion technique, *IEEE J. Solid-State Circuits,* **SC-19**: 828–836, 1984.
8. M. J. Demler, *High-Speed Analog-to-Digital Conversion,* New York: Academic Press, 1991.
9. T. Kumamoto et al., An 8-bit high-speed CMOS A/D converter, *IEEE J. Solid-State Circuits,* **SC-21**: 976–982, 1986.
10. A. G. F. Dingwall and V. Zazzu, An 8-MHz CMOS subranging 8-bit A/D converter, *IEEE J. Solid State-Circuits,* **SC-20**: 1138–1143, 1985.
11. B. Song, S. Lee, and M. F. Tompsett, A 10-b 15-MHz CMOS recycling two-step A/D converter, *IEEE J. Solid-State Circuits,* **25**: 1328–1338, 1990.
12. B. Razavi and B. A. Wooley, A 12-b 5-Msample/s two-step CMOS A/D converter, *IEEE J. Solid-State Circuits,* **27**: 1667–1678, 1992.
13. S. H. Lewis and P. R. Gray, A pipelined 5-Msample/s 9-bit analog-to-digital converter, *IEEE J. Solid-State Circuits,* **SC-22**: 954–961, 1987.
14. B. Song, M. F. Tompset, and K. R. Lakshmikumar, A 12-bit 1-Msample/s capacitor error-averaging pipelined A/D converter, *IEEE J. Solid-State Circuits,* **23**: 1324–1333, 1988.
15. W. C. Black, Jr. and D. A. Hodges, Time interleaved converter arrays, *IEEE J. Solid-State Circuits,* **SC-15**: 1022–1029, 1980.
16. C. S. G. Conroy, D. W. Cline, and P. R. Gray, An 8-b 85-MS/s parallel pipeline A/D converter in 1μm CMOS, *IEEE J. Solid-State Circuits,* **28**: 447–454, 1993.
17. J. Franca, A. Petraglia, and S. Mitra, Multirate analog-digital systems for signal processing and conversion, *Proc. IEEE,* **85**: 242–262, 1997.
18. S. Norsworthy, R. Schreier, and G. Temes, *Delta-Sigma Data Converters, Theory, Design and Simulation,* New York: IEEE Press, 1997.
19. J. C. Candy, A use of double integration in sigma delta modulation, *IEEE Trans. Commun.,* **COM-33**: 249–258, 1985.
20. J. C. Candy and O. J. Benjamin, The structure of quantization noise from sigma-delta modulation, *IEEE Trans. Commun.,* **COM-29**: 1315–1323, 1981.
21. B. E. Boser and B. A. Wooley, The design of sigma-delta modulation analog-to-digital converters, *IEEE J. Solid-State Circuits,* **23**: 1298–1308, 1988.
22. K. S. Chao et al., A higher order topology for interpolative modulators for oversampling A/D converters, *IEEE Trans. Circuits Syst.,* **37**: 309–318, 1990.
23. D. R. Welland et al., A stereo 16-bit delta-sigma A/D converter for digital audio, *J. Audio Eng. Soc.,* **37** (6): 476–486, 1989.
24. S. Mohsen Moussavi and B. H. Leung, High-order single-stage single-bit oversampling A/D converter stabilized with local feedback loops, *IEEE Trans. Circuits Syst.-II: Analog Digital Signal Process.,* **41**: 19–25, 1994.
25. Y. Matsuya et al., A 16-bit oversampling A-to-D conversion technology using triple-integration noise shaping, *IEEE J. Solid-State Circuits,* **SC 22**: 921–929, 1987.
26. M. Rebeschini et al., A 16-bit 160-kHz CMOS A/D converter using sigma-delta modulation, *IEEE J. Solid-State Circuits,* **25**: 431–440, 1990.
27. B. P. Brandt and B. A. Wooley, A 50-MHz multibit sigma-delta modulator for 12-b 2-MHz A/D conversion, *IEEE J. Solid-State Circuits,* **26**: 1746–1756, 1991.
28. G. Yin and W. Sansen, A high-frequency and high-resolution fourth-order $\Sigma\Delta$ A/D converter in BiCMOS technology, *IEEE J. Solid-State Circuits,* **29**: 857–865, 1994.
29. M. Sarhang-Nejad and G. C. Temes, A high-resolution multibit $\Sigma\Delta$ ADC with digital correction and relaxed amplifier requirements, *IEEE J. Solid-State Circuits,* **28**: 648–660, 1993.
30. B. H. Leung and S. Sutarja, Multibit Σ-Δ A/D converter incorporating a novel class of dynamic element matching techniques, *IEEE Trans. Circuits Syst.-II: Analog Digital Signal Process.,* **39**: 35–51, 1992.
31. F. Chen and B. H. Leung, A high resolution multibit sigma-delta modulator with individual level averaging, *IEEE J. Solid-State Circuits,* **30**: 453–460, 1995.
32. R. T. Baird and T. S. Fiez, Linearity enhancement of multibit $\Delta\Sigma$ A/D and D/A converters using data weighted averaging, *IEEE Trans. Circuits Syst.-II: Analog Digital Signal Process.,* **42**: 753–762, 1995.
33. S. Chu and S. Burrus, Multirate filter designs using comb filters, *IEEE Trans. Circuits Syst.,* **CAS-31**: 913–924, 1984.
34. E. Dijkstra et al., On the use of modulo arithmetic comb filters in sigma delta modulators, *Int. Conf. Acoustics, Speech, Signal Process.,* New York, April 11–14, 1988, pp. 2001–2004.
35. B. P. Brandt and B. Wooley, A low-power, area-efficient digital filter for decimation and interpolation, *IEEE J. Solid-State Circuits,* **29**: 679–687, 1994.

JOSÉ EPIFÂNIO DA FRANCA
CARLOS AZEREDO LEME
JOÃO CALADO VITAL
Instituto Superior Técnico

ANALYSIS AND DESIGN ALGORITHMS FOR CONTROL SYSTEMS. See SOFTWARE FOR CONTROL SYSTEM ANALYSIS AND DESIGN, SINGULAR VALUE DECOMPOSITION.

ANALYSIS, CEPSTRAL OF SPEECH. See CEPSTRAL ANALYSIS OF SPEECH.

ANALYSIS, COST. See COST ANALYSIS.

ANALYSIS, COST-BENEFIT. See COST-BENEFIT ANALYSIS.

ANALYSIS, INTERVAL, FOR CIRCUITS. See INTERVAL ANALYSIS FOR CIRCUITS.

ANALYSIS, NETWORK. See NETWORK ANALYSIS USING LINEARIZATION.

ANALYSIS, OF CIRCUITS IN THE FREQUENCY DOMAIN. See FREQUENCY-DOMAIN CIRCUIT ANALYSIS.

ANALYSIS OF FAILURE MODES AND EFFECTS. See FAILURE MODES AND EFFECTS ANALYSIS.

ANALYSIS OF MICROWAVE AND MILLIMETER-WAVE STRUCTURES. See SPECTRAL-DOMAIN ANALYSIS.

ANALYSIS OF RELIABILITY DATA. See STATISTICAL ANALYSIS OF RELIABILITY DATA.

ANALYSIS, QUALITATIVE OR DYNAMIC CIRCUITS. See QUALITATIVE ANALYSIS OF DYNAMIC CIRCUITS.

ANALYSIS, SENSITIVITY. See SENSITIVITY ANALYSIS.

ANALYSIS, SPEECH. See SPEECH ANALYSIS; SPEECH PROCESSING

ANALYSIS, SYSTEMS. See SYSTEMS ANALYSIS.

ANALYSIS, TRANSIENT. See TRANSIENT ANALYSIS.

ANALYZER, POWER SYSTEM. See POWER SYSTEM MEASUREMENT.

ANIMATION, COMPUTER. See COMPUTER ANIMATION.

ANNUNCIATORS. See ALARM SYSTEMS.