SAMPLE-AND-HOLD CIRCUITS

Sample-and-hold circuits were first introduced as a front end for analog-to-digital data converters. Processing electrical information in a discrete or digital fashion appears to be more reliable, repeatable, and accurate than in the analog domain. However, converting analog signals to digital ones can be often disrupted if the input signal changes during the conversion cycle. The exact moment in time when the input is sensed and compared with a reference can be different across the data converter, resulting in aperture errors. Consequently, it appears useful to memorize the input signal and hold it constant for the comparators that perform the conversion to the digital domain. Among different data-converter architectures, only the so-called "flash" ones can perform without requiring a sample-and-hold circuit although its usage is usually beneficial even in this case.

Another application for sample-and-hold circuits is as a back end for digital-to-analog data converters. The analog voltage generated by these converters is subject to glitches due to the transitions occurring between consecutive digital input codes. A sample-and-hold circuit can be used to sample the analog voltage between the glitches, effectively smoothing the output waveform between two held output-voltage levels. Then, a low-pass filter following the sample-and-hold circuit is able to restore the continuous-time analog waveform much more efficiently than in the presence of glitches.

Switched-capacitor or switched-current signal processing inherently requires the input signal to be sampled and held for subsequent operations. There is a widespread usage of this type of processing; hence a variety of applications employ some sort of sample-and-hold circuit as a front end. It can be inferred that any sampled-data system requires a sample-and-hold operation at some point.

When used inside a larger system, the performance of a sample-and-hold circuit could limit the overall performance. Speed, accuracy, and power consumption are a few criteria to be observed during the design process or simply from a user's perspective. This article is intended to describe and present different implementations of sample-and-hold circuits as well as associated nonidealities. The following section includes a list of errors that make a real-life implementation different from the ideal sample-and-hold circuit model. The third section describes different implementations of these circuits organized into metal oxide semiconductor (MOS) transistor-based open-loop architectures, MOS-transistor-based closed-loop architectures, bipolar-device-based architectures, and current-mode architectures. The last section of the article outlines some conclusions and provides a brief overview regarding modern applications of sample-and-hold circuits.

SAMPLE-AND-HOLD PERFORMANCE SPECIFICATIONS

A simplified model of a sample-and-hold circuit is a switch connected between the input and one terminal of a holding capacitor. The other terminal of the capacitor is tied to a reference voltage (ground), and the output of the sample-and-hold is the voltage across the capacitor. The switch is controlled by a digital signal that determines the sampling time and the hold duration. Ideally, the output voltage tracks the input exactly when the switch is closed (track or sample mode) and stores a sample of the input voltage when the

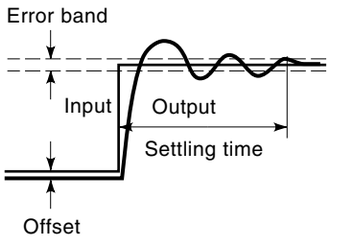

Figure 1. Offset and settling time during sample mode. The output of a sample-and-hold circuit settles to a step at the input.

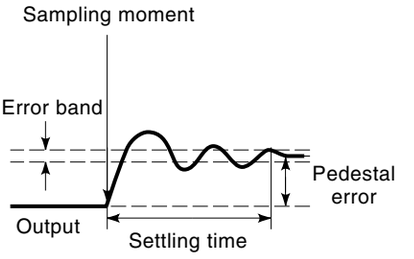

Figure 3. Settling time and pedestal voltage during the sample-to-hold transition.

switch opens (hold mode). In reality the behavior of the circuit deviates from this ideal model affecting the performance. The following subsections describe the specifications used to characterize sample-and-hold circuits.

Sample-Mode Specifications

The operation during the sample mode is similar to a voltage or current amplifier. Thus any specifications used to characterize an amplifier can be used to characterize the sample-and-hold circuit in sample mode. Some of the key specifications in this mode of operation are offset, settling time, gain error, and nonlinearity. The offset is defined as the difference between the input and the output with no signal applied at the input. This is shown in Fig. 1. The time it takes for the output to settle within a certain error band around its final value with a step applied at the input is called the settling time. The size of the step is usually full scale as shown in Fig. 1 unless specified otherwise. Gain error and nonlinearity are both steady-state errors that describe the deviation of the magnitude transfer characteristic from a straight line with a slope of 1. As represented in Fig. 2, the gain error appears as the deviation of the slope of line P_1P_2 from 45° and can be expressed as $G_{error} = 1 - \tan\theta$. One definition of nonlinearity is the maximum deviation of the transfer characteristic from line P_1P_2. This is also shown in Fig. 2. Other sample-mode specifications include bandwidth, slew rate, distortion, and noise, which also characterize a general amplifier, are defined in a similar fashion.

Sample- to Hold-Mode Specifications

The switching from sample to hold mode is accompanied by transients that appear in the output. The time it takes for these transients to settle within a given error bound is called the settling time, similar to the corresponding sample-mode specification discussed in the preceding subsection. This nonideal behavior is shown in Fig. 3. Another nonideal effect is also presented in Fig. 3 and is called the pedestal error. This error appears as an offset in the output signal and is due to charge injected from the sampling switch, which could be dependent on the input signal, implying nonlinearity.

Aperture time is another important specification and is defined as the amount of time it takes for the sampling switch to open. This is attributed to the finite slope of the clock and a gradual transition of the sampling switch from a low-impedance (closed) to a high-impedance (opened) state. The signal stored at the output is a function of the aperture delay t_a, the analog delay t_{da} of the signal through the sample-and-hold circuit, and the digital delay t_{dd} of the clock signal that turns off the switch. It can be ultimately characterized as shown in Fig. 4 by an effective aperture time t_{eff}. The analog delay is caused by the frequency response of the sample-and-hold circuit, and the digital delay is produced by any logic in the path of the external clock to the switch. The aperture jitter is the variation in the hold signal delay. In some applications, such as analog-to-digital data conversion, a constant delay in the sampling time is not important. However, aperture jitter could be extremely damaging, adding significant noise to the output signal and effectively lowering the resolution of the system.

Hold-Mode Specifications

Ideally, the holding capacitor should be completely isolated from any interference during the hold mode and the output signal should remain constant. In a real-life implementation there is always a small current flowing across the holding ca-

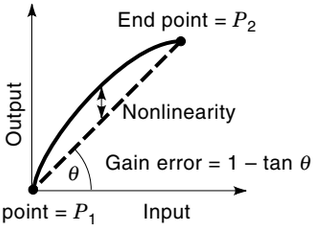

Figure 2. Gain error and nonlinearity. A sample-and-hold circuit exhibits gain error as deviation of line P_1P_2 from a 45° slope, and nonlinearity as curvature in the input–output characteristic.

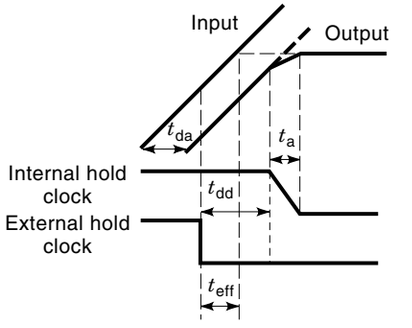

Figure 4. Time delays that determine the effective aperture time during the sample-to-hold transition: t_a, aperture delay; t_{dd}, digital delay; t_{da}, analog delay; t_{eff}, effective aperture time.

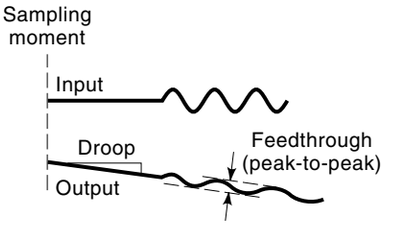

Figure 5. Droop and feed-through errors during hold mode.

pacitor during the hold mode. This nonideal effect produces droop in the output signal as shown in Fig. 5. As an example, in an analog-to-digital data converter the droop should not exceed one-half of the least-significant bit value during the conversion cycle. Generally, the droop current could be a bipolar base current or simply junction leakage. Another specification during the hold mode is the feedthrough from the input signal. This occurs through parasitic capacitance between the input and the output nodes of a sample-and-hold circuit, although it can occur between other nodes in the system and the output. Figure 5 also presents this effect. Again, in reference to the data-converter example, the feedthrough peak-to-peak amplitude should not exceed one least significant bit of the converter.

There are a few other factors that influence the performance during the hold mode. Nonidealities in the holding capacitor can cause dielectric absorption. Consequently, the capacitor exhibits a memorylike behavior, corrupting the newly stored sample in the direction of the previous sample. Electrical noise is also a factor, and the value of the capacitor plays an important role since the kT/C term (k is the Boltzmann constant, T the absolute temperature in kelvin, and C the holding capacitor value) is usually the main contributor.

Hold- to Sample-Mode Specifications

When switching from the hold to sample mode, there is a delay between the sampling edge of the clock and the moment the output settles to the steady state when it tracks the input. This time is known as the acquisition time and is also determined by an error band. Since the input can change significantly while the circuit is in hold mode, the acquisition time is usually specified for a full-scale change in the input unless mentioned otherwise.

IMPLEMENTATION OF SAMPLE-AND-HOLD CIRCUITS

If one starts with the basic ideal model for a sample-and-hold circuit, there are various ways to implement a switch and a capacitor in either discrete or monolithic forms. The interface between this model and the input signal or adjacent system blocks implies the usage of additional circuitry. There is also added complexity from dealing with different circuit nonidealities, an increase in operating speed, or a decrease in power consumption. The following subsections embrace this gradual approach from a basic model to a more complex structure, which provides an implementation that is close to real life. A variety of topologies and implementations is presented including open- and closed-loop, bipolar, and MOS-based, as well as current-mode circuits.

Open-Loop Sample-and-Hold Circuits

The simplest voltage-mode sample-and-hold circuit requires only two elements: a switch and a capacitor, as shown in Fig. 6 (1–3). The switch is controlled by a clock signal ϕ that is turned on and off each sample period. When the switch is on, the input voltage appears on the capacitor, and the circuit is actually in track mode. When the switch is turned off, the signal voltage at that instance is sampled on the capacitor, which holds the voltage constant until the next track phase. Although this ideally implements the desired sample-and-hold function using only two elements, the difficulty in realizing an ideal switch severely limits the performance of such a design. Real switches are implemented using MOS or bipolar transistors or diode bridges, each of which has its own idiosyncrasies. This subsection will primarily focus on the use of MOS switches, while circuits utilizing bipolar transistors or diodes will be discussed in the following subsection.

A MOS transistor deviates from an ideal switch in several ways, the most obvious of which is in terms of on-resistance. When the MOS switch is turned on, it typically will have a low drain-to-source voltage (V_{DS}) and a high gate-to-source voltage (V_{GS}) [n-type MOS assuming (NMOS), opposite polarity for p-type MOS (PMOS)], causing it to operate in the triode or nonsaturation region, where the drain-to-source current (I_{DS}) is given by (4)

$$I_{DS} = K' \frac{W}{L}[(V_{GS} - V_T)V_{DS} - \tfrac{1}{2}V_{DS}^2] \tag{1}$$

which, when $V_{DS} \ll (V_{GS} - V_T)$, can be approximated as

$$I_{DS} \approx K' \frac{W}{L}(V_{GS} - V_T)V_{DS} \tag{2}$$

where V_T is the threshold voltage. The large-signal on-resistance of the switch is then given by the voltage across the device, V_{DS}, divided by the current through it, I_{DS}, yielding

$$R_{switch} \approx \left(K' \frac{W}{L}(V_{GS} - V_T)\right)^{-1} \tag{3}$$

The switch on-resistance is clearly nonzero, meaning that the circuit will exhibit some transient exponential settling when the switch first turns on and then a low-pass filtering response during steady-state operation while it remains on (for input signal frequencies greater than dc). Clearly the switch resistance must be designed properly to provide sufficient levels of settling and bandwidth for the speed and precision needed.

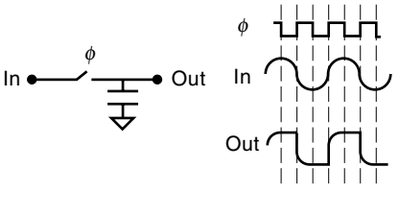

Figure 6. Simple sample-and-hold circuit with example signals shown. When ϕ is high, the output tracks the input and when ϕ is low, the output is held.

Not only is the resistance of the switch nonzero and dependent on the semiconductor process parameter K' and device size (W and L), but it also varies with the voltage on the source of the transistor, which will be a function of the input voltage. If, as is done in the simplest cases, the gate of the transistor is tied to a constant-supply-voltage level while the switch is on, then the switch resistance varies with the input signal, causing nonlinear dynamic circuit operation. If an NMOS switch is used, then its on-resistance will rise toward infinity as the input voltage approaches $V_{DD} - V_{T,n}$, limiting the input range to voltages well below this. Similarly, a PMOS switch will limit the input to voltages well above $V_{SS} + |V_{T,p}|$. A wider input range is obtained by using a complementary MOS (CMOS) switch, consisting of an NMOS and a PMOS switch in parallel, with the PMOS device driven by the inverted version of the sampling signal ϕ. This configuration enables sampling of an input range extending from V_{SS} to V_{DD}, although on-resistance of the switch will still vary significantly over the entire range, giving rise to nonlinear dynamic performance. This effect can be avoided by driving the switch gates with a bootstrapped version of the input signal while the switches are on, such that both gates and sources will vary similarly with respect to the input (5). This removes the input signal dependence from the V_{GS} term in Eq. (3), making the on-resistance independent of the input, at least to first order. In reality, higher-order effects will eventually come into play to limit the linearity, but this technique is useful in gaining some performance improvement.

Another limitation of the MOS switch is charge injection and clock feedthrough, which are sometimes used interchangeably to describe two effects that occur when the switch is turned off. When the transistor is on, charge collects under the gate of the transistor to form the channel from drain to source. When the switch is turned off, this charge exits the channel primarily to the drain and source, with the proportional split depending on speed of the clock transition and the impedance seen by the charge in each direction (6). The part of the charge that exits toward the sampling capacitor will cause the capacitor voltage to drop slightly (assuming an NMOS transistor with negative channel charge) from its immediately previous value. This would only create a constant pedestal error in the sample-and-hold circuit output if the packet of charge were always the same, but unfortunately this is not the case. The channel charge is approximately given by

$$Q_{\text{channel}} \approx C_{ox} WL(V_{GS} - V_T) \tag{4}$$

which reveals that, if the gate of the switch is tied to a constant supply voltage while turned on, then the channel charge will be signal dependent. In reality, this signal dependence of the injected charge is a major source of nonlinearity in the sample-and-hold circuit.

The second effect, often termed *clock feedthrough,* is caused by the MOS overlap capacitance between the gate and source or drain connected to the sampling capacitor. As the gate voltage is dropping from a high on-voltage to a low off-voltage, the transistor actually shuts off when the gate is approximately a threshold voltage above the source or drain voltage. As the gate voltage continues to fall further, the voltage step is capacitively coupled onto the sampling capacitor through the MOS overlap capacitance, causing the voltage on the sampling capacitor to change. Both of these charge effects can be reduced by driving the gate with a bootstrapped version of the input signal, which makes the effects independent of signal to first order, just as in the case of switch on-resistance. The use of a full CMOS switch rather than only a single NMOS or a single PMOS device also gives a very rough cancellation of these effects. The NMOS device will inject negative channel charge onto the sampling capacitor, while the PMOS switch will inject positive channel charge. Similarly, the charge coupled through the overlap capacitances will also be in opposing directions. Unfortunately, these cancellations do not hold as the input is varied over its input range and so give little benefit.

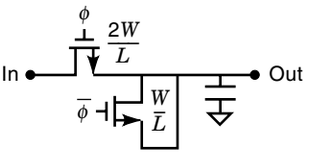

Figure 7. Simple sample-and-hold circuit with dummy-switch compensation for reducing the charge injection. The dummy device is half the size (W/L vs. $2W/L$) of the series switch device.

Dummy-switch cancellation is another technique that can be used, though generally with limited effectiveness, to reduce charge injection and clock feedthrough. As shown in Fig. 7, a dummy-switch device with half the gate area of the main switch is placed on the capacitor side of the sampling switch and is clocked with an inverted version of the sampling signal ϕ. Thus, during acquisition, the main switch is on and the dummy switch is off. At the sampling instant, the dummy switch is turned on, causing it to pull charge from the sampling capacitor to form the channel under its gate. This charge approximately cancels that injected by the main switch, assuming roughly half of the main switch charge was injected toward the capacitor.

The simple sample-and-hold circuit using a MOS switch can also experience input signal feedthrough during the hold mode due to the MOS overlap capacitances, especially when sampling high-speed inputs. When the circuit is in hold mode, the MOS gate will be pulled to the low supply voltage (assuming an NMOS switch), typically by a digital gate's output. The digital gate will have a finite, nonzero resistance from its output to the low supply voltage, yielding an effective circuit such as that shown in Fig. 8. Analysis of this circuit reveals a nonzero transfer function from the input to the sampled output node, causing the feedthrough. Care must be taken during the design process to characterize the level of feedthrough that occurs and keep it sufficiently low through a low-resistance gate drive or more elaborate design modifications.

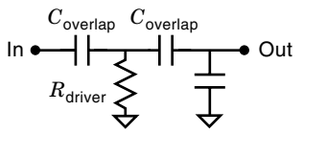

Figure 8. Effective circuit of the simple sample-and-hold circuit using an NMOS switch while in hold mode. The gate is driven low by a circuit with nonzero output resistance, such as a digital gate or clock driver. The MOS overlap capacitances couple the input signal to the output even with the MOS switch turned off.

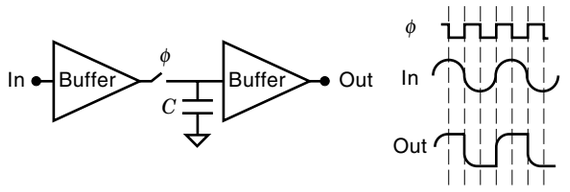

Figure 9. Simple sample-and-hold circuit with input and output buffering.

A fundamental trade-off occurs in this simple sample-and-hold circuit between speed and linearity. In order for the circuit to have high bandwidth and fast settling, a large width transistor with low on-resistance is needed. However, larger width also means more channel charge, which then increases the charge injection and generally reduces the circuit's linearity performance. While the sampling switch must be designed to meet the speed and settling requirements of the given application, it should not be unnecessarily overdesigned to have excessively low on-resistance.

Yet another limitation of the design shown in Fig. 6 is its lack of drive capability while in hold mode. If this circuit is loaded by a resistive or switched capacitive load, then the sampled signal voltage on the capacitor will leak off while the sampling switch is off. This problem is easily addressed by adding a unity-gain buffer at the output. The loading that this sample-and-hold circuit places on the input signal source can also be a problem in some applications, particularly for sources with large output resistance or limited slew rate. This is also easily remedied with another unity-gain buffer, now placed at the input. The resulting design is shown in Fig. 9. While these buffers do solve the loading problems, in reality each will have some amount of dc offset, gain error (i.e., the gain of the buffer is not exactly unity), and nonlinearity, all of which will directly affect the offset, gain error, and nonlinearity of the overall sample-and-hold circuit. The dynamic performance of the buffers is also a concern. When the sampling switch is turned on, the input buffer may have to supply a sudden impulse of current to the sampling capacitor, typically causing a transient settling response at the buffer output. This may include both slewing and linear settling behavior, both of which must settle out to the desired level of precision before the sampling switch is opened. The buffers will also have limited bandwidth, creating a steady-state amplitude and phase error that varies with input frequency, independent of whether the buffer transient responses have settled or not.

While the simple open-loop sample-and-hold circuit shown in Fig. 6 does have many performance limitations, some of these can be reduced or avoided by modifying the circuit architecture. An alternative open-loop sampling using three clocks is shown in Fig. 10. While ϕ_1 and ϕ_{1d} are high, the input voltage is sampled across the capacitor through the series resistances of the two switches. The actual sampling instant occurs when ϕ_1 falls, open-circuiting the right side of the capacitor. ϕ_{1d} then falls immediately afterward, disconnecting the capacitor from the input signal. When ϕ_2 goes high, an inverted version of the sampled input voltage appears at the right side of the capacitor and is buffered to the output. Note that this circuit cannot be termed a track-and-hold circuit, since the output never tracks the input directly, but instead has its input grounded while the input capacitor is being charged. The advantage of this design is its avoidance of signal-dependent charge-injection effects. The ϕ_{1d} falling edge is intentionally delayed from the ϕ_1 falling edge in order to make only the ϕ_1 switch's charge injection important (otherwise ϕ_{1d} is the same as ϕ_1). After the ϕ_1 switch is off, any further charge injection or transients at the left-hand side of the capacitor will not affect the charge that has been isolated on the node at the right-hand side of the capacitor. In addition, because the ϕ_1 switch always operates at ground potential, its channel charge (and on-resistance) will always be the same, without the need for a bootstrapped gate drive. The only signal dependence left in the charge injected by the ϕ_1 switch is due to variations in the impedance looking back toward the input signal source as the input voltage varies.

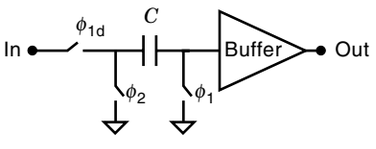

Figure 10. Modified sample-and-hold circuit using additional switches and an output buffer.

While this sample-and-hold circuit has an advantage with respect to charge injection, it also has a disadvantage in that it is sensitive to parasitic capacitance at the right-hand side of the sampling capacitor. Parasitic capacitance to ground at this point will lead to charge redistribution during ϕ_2, thereby attenuating the sampled voltage and causing a gain error in the system. In addition, if the parasitic capacitor happens to be nonlinear, such as the reverse-biased diode junction capacitance associated with the ϕ_1 switch, then nonlinearity will also be introduced along with the attenuation. These effects are reduced by making the sampling capacitor sufficiently larger than the parasitic capacitors, but then settling is affected, and another trade-off between speed and linearity can be seen. An alternative solution to these problems is to use a closed-loop architecture, which exploits the benefits of negative feedback within its design. These architectures are discussed in a following subsection.

From an ideal standpoint, one would prefer a sample-and-hold circuit that did not have any reset phases or a track phase, but would instead only transition from one hold state to the next hold state, thus giving the circuitry that follows the maximum time to utilize the held signal. Although the circuits shown in Figs. 9 and 10 provide a valid output for approximately half of the sampling period, this can be extended to a full sampling period simply by using two or more parallel sample-and-hold circuits that operate on different clock phases and multiplexing among their outputs, as shown in Fig. 11 (7). This multirate polyphase design approach trades off power and circuit area in exchange for an increase in effective sampling rate and a maximized valid output time. While this works perfectly in theory, practical issues such as matching from channel to channel will limit its range of application. For example, if four parallel channels are used, each with a different dc offset and different gain, then the single output after multiplexing will contain pattern noise even when a zero input is applied. In addition, a modulation effect can occur between the pattern noise and the signal, mixing high-frequency components to low frequency and vice versa. If the input is band-limited and sufficiently oversampled,

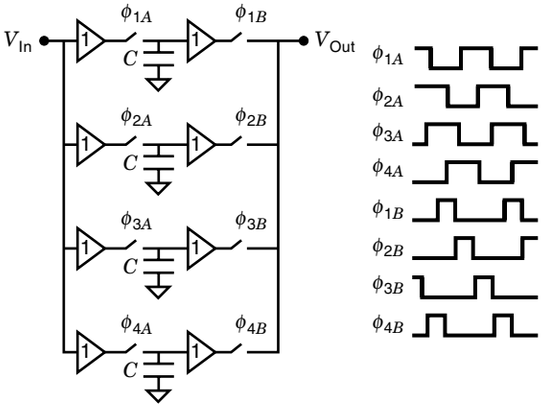

Figure 11. Polyphase sample-and-hold architecture demonstrated using four parallel channels, each phased separately, to achieve a fourfold increase in effective sampling rate.

then the pattern noise and modulation artifacts can be kept spectrally separate from the signal component and thus filtered out in a later stage.

While all of the designs shown thus far have used a single input voltage relative to ground and have provided a single output also relative to ground, each of the circuits can easily be converted to a fully differential configuration. Figure 12 demonstrates a fully differential version of the circuit originally shown in Fig. 10, in which the input is applied as the difference between two node voltages, and the output is supplied as the difference between the two output node voltages. Fully differential configurations are widely used in high-performance analog designs due to a variety of advantages. If one assumes that corresponding switches and components are matched, a fully differential circuit will provide cancellation of dc offsets associated with charge injection and clock feedthrough due to the (ideally) same charge being injected onto both sampling capacitors. Because the output is used differentially, only the difference between the sampling capacitor voltages is important, and so the identical charge that was injected into both capacitors cancels out. Fully differential circuits also have low (ideally zero) even-order harmonic distortion components due to their symmetry if both sides of the differential circuitry are matched. These designs also exhibit higher power-supply rejection than single-ended designs due to power-supply noise, causing a similar response on both sides of the differential circuit, which then cancels out when only the difference in the two output nodes is considered. In addition, even though the voltage swing of the input and output nodes remains the same, the fully differential circuit will

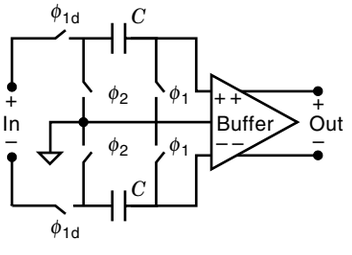

Figure 12. Fully differential version of the sample-and-hold circuit shown in Fig. 10.

have double the input and output swing as the single-ended design, due to the fact that the signals are now $V_{plus} - V_{minus}$ rather than $V_{plus} - V_{ground}$. This helps keep the signal higher above the noise floor, thus maintaining a higher signal-to-noise ratio. While fully differential circuits provide numerous advantages, they also require approximately double the area and power of a single-ended design, and extra circuitry and complexity are needed to control the common-mode voltages of nodes throughout the design. However, most designers of high-performance analog circuitry have accepted the need for differential circuitry and utilize it extensively.

SAMPLE-AND-HOLD CIRCUITS USING BIPOLAR AND DIODE DEVICES

The basic components of sample-and-hold circuits can be also implemented using bipolar technology. Semiconductor diodes and bipolar transistors are potential candidates for performing the switch function as well as input and output buffering. The holding capacitor could be a metal-metal or a MOS structure in the case of monolithic implementations or just a discrete component otherwise. The first part of this subsection will discuss sample-and-hold circuits in which the required switching operation is implemented using bipolar transistors. Following this, diode-bridge sample-and-hold circuits are introduced and the subsection is concluded with a discussion regarding performance comparison between MOS and bipolar implementations.

Although the bipolar transistor exhibits similar behavior to a MOS transistor, when used in amplifier configurations, there are fundamental differences that prohibit its usage directly as a switch in the same fashion as a MOS-based one. As an example, for an n–p–n bipolar transistor, when the base voltage is lower or equal to its collector and emitter voltages, the resistance seen between the emitter and collector is high. When its base voltage is higher (both base–collector and base–emitter junctions are forward biased), the resistance "seen" between the emitter and collector becomes low; however, it is very nonlinear. Consequently, an arrangement such as that in Fig. 6 from the previous subsection, where the switch was replaced with an n–p–n transistor, is not possible. Consequently, bipolar transistors can be used as switches although not in series with the holding capacitor. An alternative is to use bipolar transistors configured as source followers in the signal path and employ bipolar-based switches to control their bias current.

Figure 13 presents a simple sample-and-hold circuit in which the switch is constituted by the combination of bipolar transistors T1–T3 and a current source I. When the clock signal is high, the circuit is in track mode. Transistors T2 and T3 can be regarded, respectively, as a closed and an open switch. The current source sinks a bias current I through T1, which behaves like an emitter follower, and the output voltage V_{out} tracks the input V_{in}. The buffer is assumed to have unity gain. When the clock signal is low, transistor T2 becomes an opened switch and transistor T3 appears as a closed one, and the current I flows entirely through transistor T3. The value of the voltage V_B (voltage at the buffer output) during this state is determined by the buffer circuitry and the value of current I, and is designed to be low enough such that transistor T1 is turned off. This is the hold state, when V_{out} is preserved on C_{hold} until transistor T1 is reconfigured again as

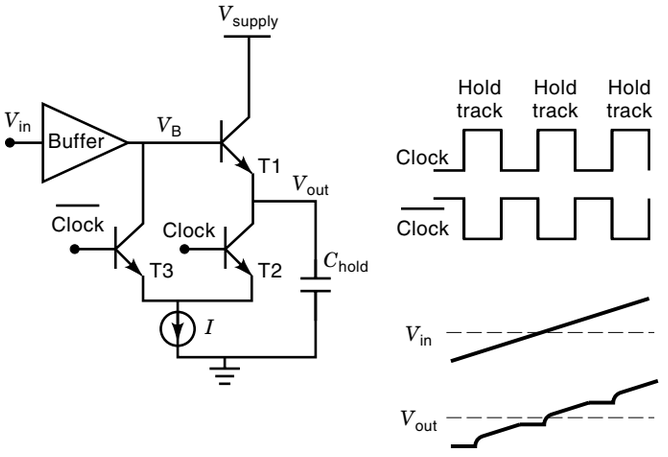

Figure 13. Sample-and-hold circuit using bipolar transistors as switches. Transistors T2 and T3 act like switches while transistor T1 and the buffer condition the signal from input to output.

an emitter follower. An example for the buffer circuit is shown in Fig. 14. In track mode, the gain of the buffer is

$$G_B = \frac{g_{m,TB}R_C}{1 + g_{m,TB}R_E} \approx 1 \tag{5}$$

if $R_C = R_E$ and $g_{m,TB} \gg 1$, where TB, R_E, and R_C are the transistor ($g_{m,TB}$ is its transconductance) and respectively the resistors from Fig. 14. The latter assumption is justified if the transconductance value is large enough, or equivalently, there is enough bias current flowing through transistor TB. Unlike the MOS transistor case, the transconductance of a bipolar device biased in the linear region is not dependent on the process parameters and is determined solely by the bias current, in a first-order approximation. This is advantageous since the equivalent R_{switch} value in the track mode depends on the source-follower transconductance value. The relation between input and output voltages of the sample-and-hold circuit from Fig. 13 can be expressed in the s-transform domain as

$$\frac{V_{out}(s)}{V_{in}(s)} = G_B \frac{g_{m,T1}}{g_{m,T1} + sC_{hold}} \tag{6}$$

where $g_{m,T1} = I/V_T$ is the transconductance value of transistor T1. As mentioned earlier, this value is dependent solely on the value of the current source I, since V_T is the thermal voltage ($V_T = kT/q \approx 26$ mV at 300 K, where k is Boltzmann's constant, q is the electron charge, and T is the absolute temperature). In the ideal case of an "on" switch in series with the holding capacitor (as in Fig. 6), the relation between the input and the output voltages is

$$\frac{V_{out}(s)}{V_{in}(s)} = \frac{1}{1 + sC_{hold}R_{switch}} \tag{7}$$

From Eqs. (6) and (7) it can be inferred that the equivalent value of the "on" switch resistor for the circuit from Fig. 13 is

$$R_{switch} = \frac{1}{g_{m,T1}} = \frac{V_T}{I} \tag{8}$$

As discussed in the open-loop sample-and-hold circuit subsection, the nonlinearity of this equivalent R_{switch} resistor is a major contributor to the degradation in performance of the sample-and-hold circuit. Another contributor is the gain error associated with the equivalent switch (source follower T1 on Fig. 13) or the input and output buffering operations. Equations (6) and (8) suggest that in a first-order approximation, there is no nonlinearity associated with transistor T1 and the only dc gain error is generated by G_B. However, the current source that generates the current I does have finite output impedance, and this introduces gain and linearity error. For example, if R_I is defined as the current-source output impedance, the dc gain of the circuit in track mode acquires a new term $g_{m,T1}R_I/(1 + g_{m,T1}R_I)$, where $g_{m,T1}$ is the transconductance of transistor T1. Also, the equivalent "on" switch resistance becomes

$$R_{switch} = \frac{R_I}{1 + g_{m,T1}R_I} \tag{9}$$

A given linearity requirement will constrain the input signal range such that there are small enough variations in the bias current I due to changes in the input.

There is no buffering at the output shown in Fig. 13, although the circuit would require it for driving a low impedance. Besides gain error, the output buffering introduces droop errors when implemented with bipolar transistors. The base current has significantly larger values than the gate current associated with MOS-transistor-based buffers. This effect can be alleviated by using differential topologies in which the base current produces a shift in the common mode but leaves the differential voltage across holding capacitors unaffected. Charge injection in bipolar implementations of switches is less destructive than that in the MOS transistor case. The reason is a smaller dependency between the amount of charge dumped into the holding capacitor at the transition moment from track to hold and the input voltage. Differential topologies also attenuate this effect as in the droop error case. Ultimately, the hold-mode feed through, from the V_b voltage (in Fig. 13) to V_{out} through the base-emitter capacitance of transistor T1 can seriously affect performance. Ways to cope with this are reducing the size (area) of transistor T1 and, again, employing differential structures that allow placing cross-coupled capacitors from the input to the output as described in Ref. 8.

Another sample-and-hold circuit based on bipolar transistors is shown in Fig. 15 (9). Transistors TN1, TN2, and TN3

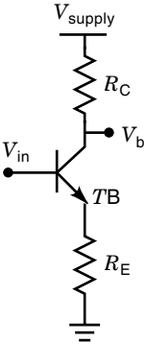

Figure 14. An input buffer for the sample-and-hold circuit in Fig. 13.

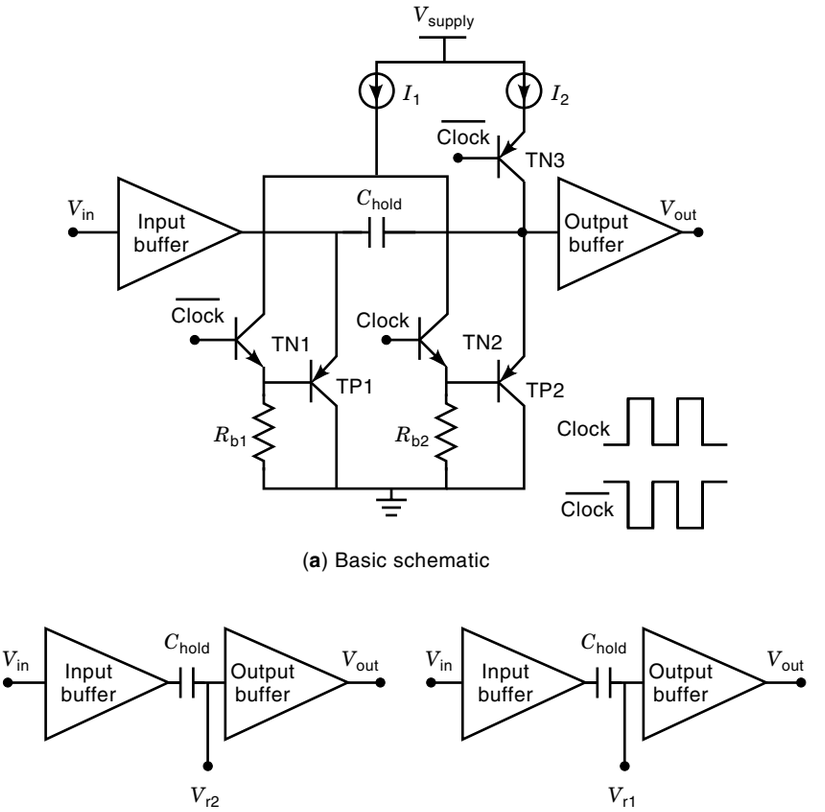

Figure 15. Sample-and-hold circuit where the holding capacitor is connected in series between input and output; (a) detailed schematic with TN1, TN2, TP1, TP2 as switches, and input–output buffering; (b) simplified schematic during track mode; (c) simplified schematic during hold mode.

are n–p–n devices that function as series switches for steering bias currents. Transistors TP1 and TP2 are p–n–p devices and work in either linear region or are turned off. For simplicity, it is assumed that the input and output buffers have unity gain. When the *clock* signal is low, transistor TN2 is turned off, and transistors TN1 and TN3 are turned on. Current I_1 [generated by the current source as shown in Fig. 15(a)] flows entirely through resistor R_{b1}, which is designed such that the voltage drop on it, IR_{b1}, turns off transistor TP1 (p–n–p type). Also, current I_2 [generated by a current source as shown in Fig. 15(a)] flows through transistor TP2 (p–n–p type), which is biased in the linear region. Removing all transistors that are turned off, the schematic can be simplified as in Fig. 15(b). The voltage V_{r2} shown on this figure can be expressed as

$$V_{r2} = V_{be,TP2} + V_{R_{b2}} = V_T \ln \frac{I_2}{I_{S,TP2}} + \frac{I_2}{\beta_{TP2}} R_{b2} \qquad (10)$$

where $V_{be,TP2}$ is the base-emitter voltage of transistor TP2, V_T is the thermal voltage, $I_{S,TP2}$ is a constant current related to transistor TP2, and β_{TP2} is the forward gain factor for transistor TP2. In other words, voltage V_{r2} is independent of the input signal and is equal to the sum of the base-emitter voltage of transistor TP2 and the voltage drop across resistor R_{b2}. The voltage on the capacitor is $V_{C_{hold}} = V_{in} - V_{r2}$, which is basically a shifted version of the input. The output voltage V_{out} is reset to V_{r2} during this track or sample phase. When the clock signal goes high transistors TN1 and TN3 turn off and TN2 turns on. Current I_1 is steered to flow entirely through resistor R_{b2}, which is designed in a similar fashion as R_{b1} such that the voltage drop on it, IR_{b2}, turns off transistor TP2. The schematic, again, can be simplified as in Fig. 15(c). The input buffer provides some current I_B, which flows through transistor TP1 biasing it in the linear region. V_{r1} can be also expressed as in Eq. (10)

$$V_{r1} = V_{be,TP1} + V_{R_{b1}} = V_T \ln \frac{I_B}{I_{S,TP1}} + \frac{I_B}{\beta_{TP1}} R_{b1} \qquad (11)$$

with similar meanings applied to transistor TP1. The output voltage becomes

$$V_{out} = V_{r1} - V_{C_{hold}}|_{sampled} = V_{r1} + V_{r2} - V_{in}|_{sampled} \qquad (12)$$

It is worth mentioning that the output voltage holds a value of the input sampled when the clock goes high, and it is reset to V_{r2} when the clock goes back to the low state. This puts an additional bandwidth requirement on the output buffer, which needs to settle more than the full range of the input signal. Also, assuming unity gain buffers, as stated initially, the output is a shifted version of the input taken with opposite sign. To achieve a perfect copy of the sampled input, a differential topology should be employed (to eliminate the shift) and either the input or output buffer should have a negative gain. Since there is no true track mode at the output, one could consider the voltage on the capacitor as a function of the input when clock is low, and derive the equivalent "on" switch resistance. However, this requires taking into account

the design of the buffers and the parasitic effects of bipolar devices, which becomes rather involved.

The most common high-speed sample-and-hold circuits employ semiconductor diode bridges. The switching times for semiconductor diodes are inversely proportional to the cutoff frequency characteristic to a given bipolar or bipolar CMOS process. Also the "on" resistance for these devices is very low compared to MOS-based switches. For a given current and reasonable sizes, the transconductance of bipolar transistors is significantly higher than that for their MOS counterparts. Since semiconductor diodes can be considered as bipolar devices with the base shorted to the collector, their "on" resistance is $R_{\text{switch}} = 1/g_{\text{m}} = V_{\text{T}}/I$, where V_{T} is the thermal voltage and I is the bias current flowing through the diode. The small "on" resistance and short switching times make the diode-bridge-based sample-and-hold circuits attractive for very-high-speed applications. Figure 16 presents such a basic sample-and-hold circuit without input and output buffering. The switch is implemented using a diode bridge and two switched current sources. When the clock is low, the switched current sources do not generate any current and present a very high impedance at their output such that the diodes turn off. This is the hold mode, when in the absence of parasitic effects such as diode leakage, feedthrough due to diode junction capacitance, or droop due to the output buffer, the output voltage preserves the sampled input voltage. When the clock goes high, current I starts flowing through the diodes and the output voltage is following the input after a short settling period. Since the resistance through the diode is fairly low, the series terminal contact resistance r_{b} of the diode becomes significant and should be taken into account. The equivalent "on" switch resistance can be expressed as

$$R_{\text{switch}} = \frac{\left(r_{\text{b1}} + \dfrac{1}{g_{\text{m1}}} + r_{\text{b2}} + \dfrac{1}{g_{\text{m2}}}\right)\left(r_{\text{b4}} + \dfrac{1}{g_{\text{m4}}} + r_{\text{b3}} + \dfrac{1}{g_{\text{m3}}}\right)}{r_{\text{b1}} + \dfrac{1}{g_{\text{m1}}} + r_{\text{b2}} + \dfrac{1}{g_{\text{m2}}} + r_{\text{b4}} + \dfrac{1}{g_{\text{m4}}} + r_{\text{b3}} + \dfrac{1}{g_{\text{m3}}}}$$
$$= r_{\text{b}} + \frac{1}{g_{\text{m}}} = r_{\text{b}} + \frac{V_{\text{T}}}{I} \tag{13}$$

if $r_{\text{b1}} = r_{\text{b2}} = r_{\text{b3}} = r_{\text{b4}} = r_{\text{b}}$ and $g_{\text{m1}} = g_{\text{m2}} = g_{\text{m3}} = g_{\text{m4}} = g_{\text{m}}$, which is equivalent to state that there are no mismatches among diodes D1–D4. Mismatches among diodes as well as

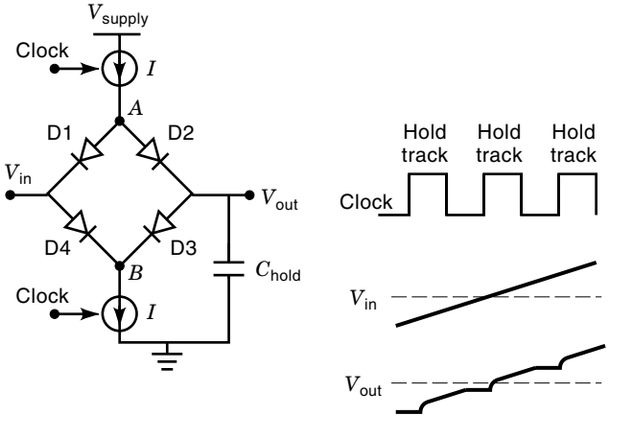

Figure 16. Basic sample-and-hold circuit using a diode bridge.

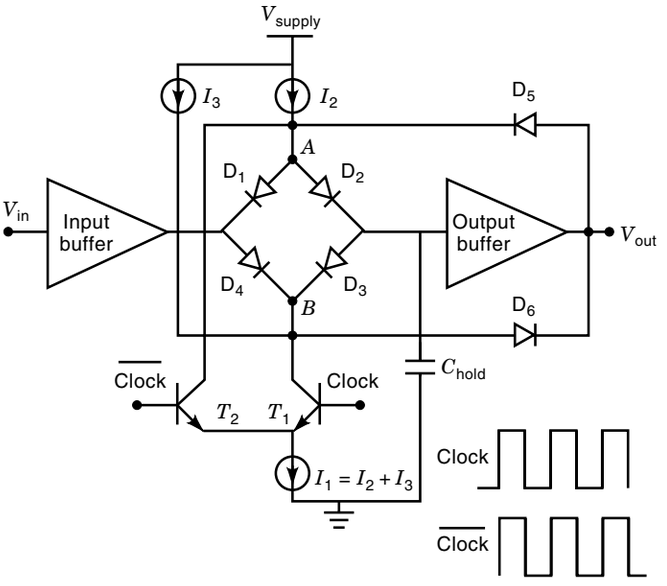

Figure 17. Diode-bridge-based sample-and-hold circuit with buffering and clamping to the output voltage.

parasitic effects mentioned earlier impact the performance of the sample-and-hold circuit from Fig. 16 by introducing nonlinearity and gain error. Fortunately, there are circuit techniques that alleviate these effects and one of them is presented later. The voltage at nodes A and B during the track mode is input dependent. However, during hold mode these voltages change to some value independent of the input, and this change couples to the output node through the diode D2 and D3 junction capacitance, creating nonlinearity. This effect can be virtually eliminated during the hold mode by clamping nodes A and B to a value dependent on the input sample. This creates a constant voltage change at these nodes from track to hold, and consequently, a constant perturbation of the held voltage, which results in offset but no nonlinearity. Figure 17 presents a circuit that performs the described clamping operation and includes input and output buffers that are assumed to have unity gain. When the clock signal is high, transistor T_2 is turned off and current I_2 flows through the diode bridge, generating a low-resistance path between the input and the holding capacitor C_{hold}. This is the track mode when the voltage on C_{hold} and the output voltage are following the input voltage. When the clock goes low, during the hold mode, transistor T_1 turns off and transistor T2 turns on, bypassing the diode bridge. Current I_3 is necessary to clamp the voltage at node B to $V_{\text{in}}|_{\text{sampled}} + V_{\text{on,D6}}$. Meanwhile, node A is clamped to $V_{\text{in}}|_{\text{sampled}} - V_{\text{on,D5}}$. So the voltage at nodes A and B changes from track to hold by $-V_{\text{on,D1}} - V_{\text{on/D5}}$ and $V_{\text{on,D3}} + V_{\text{on,D6}}$, respectively. These V_{on} voltages are dependent on the diode sizes and bias current $[V_{\text{on}} = V_{\text{T}} \ln(I_{\text{b}}/I_{\text{S}})]$ so the coupling from nodes A and B to the voltage on C_{hold} through the junction capacitance of diodes D2 and D3 is signal independent. Also, during the track mode diodes D5 and D6 are turned off since the output voltage follows the input and the voltages at node A and B are $V_{\text{in}} + V_{\text{on,D1}}$ and $V_{\text{in}} - V_{\text{on,D4}}$, respectively. It is important to note that the range of the input signal is limited to keep the diodes and current sources in their desired regions of operation.

There are various architectures (10–12) based on Fig. 17 that deal with alleviating undesired effects including those mentioned earlier as well as input and output buffer linearity and speed. Ultimately, the trade-offs made to improve the performance of the sample-and-hold circuit need to be tailored according to the overall application requirements. Bipolar-transistor- or diode-based switches appear as a viable option; however, there are a few key issues that can tip the balance between a MOS and a bipolar implementation (2). The equivalent "on" resistance is more linear and potentially smaller in the bipolar case. Also, bipolar transistors require a smaller clock voltage swing to be controlled as switches, which results in a more accurate sampling instant. On the negative side, they limit significantly the input range and generally create an offset (eventually due to mismatches) between the input and output voltages.

Closed-Loop Sample-and-Hold Circuits

Several of the problems that plague open-loop sample-and-hold circuits can be avoided through the use of closed-loop negative feedback in the design. Some of the circuits included here still perform an open-loop sampling operation, but they utilize negative feedback in the buffer stage that follows. One of the simplest configurations that is often used in switched-capacitor circuitry is the reset integrator, shown in Fig. 18 (13). During ϕ_1, the input is sampled onto C_1 and the switched-capacitor integrator is reset to zero by discharging C_F. On ϕ_2, the charge on C_1 is integrated onto C_F, yielding $V_{out} = -(C_1/C_F)V_{in}$. Note that, as discussed in the subsection concerning open-loop topologies, the switches at the virtual ground side of the sampling capacitor open slightly before those on the opposite side. The use of the switched-capacitor integrator also avoids problems associated with the parasitic capacitance at the right-hand side of the sampling capacitor to ground, since this node always returns to the same virtual ground potential during both sample and hold modes. Clearly, the design of an operational amplifier with sufficient gain and bandwidth is critical to the circuit's performance. Droop during the hold mode is typically low in this design and is signal independent, caused by the leakage current through reverse-biased source- or drain-to-bulk diodes associated with the turned-off switches at the operational amplifier input. A disadvantage of this approach is that the output is not valid until near the end of ϕ_2, so an extra clock phase is often needed after ϕ_2 (shown by ϕ_3 in Fig. 18) for the circuitry that follows this stage to utilize the output.

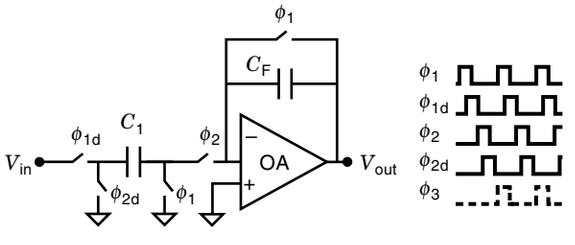

Figure 18. A reset integrator sample-and-hold circuit, which uses open-loop sampling and an operational amplifier in closed-loop feedback as an integrator.

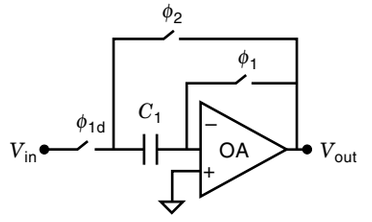

Figure 19. A sample-and-hold circuit based on switched-capacitor techniques and using a single capacitor both for sampling and in feedback around the operational amplifier.

A similar circuit that avoids the extra clock phase is shown in Fig. 19. The input is sampled onto C_1 relative to the virtual ground of the operational amplifier on ϕ_1, and then is placed in feedback around the amplifier during ϕ_2, yielding $V_{out} = V_{in}$. Thus, C_1 serves as both sampling and feedback capacitor. An advantage of this technique is that the dc offset and low-frequency noise of the amplifier (such as $1/f$ noise) is removed by this sampling scheme. The disadvantage is that the operational amplifier is required to sink the input current needed to charge C_1 initially, potentially increasing the operational amplifier power dissipation, but it still does not offer any input buffering to the driving source.

Another approach to a switched-capacitor sample-and-hold circuit is shown in Fig. 20 in fully differential form and consists of a simple first-order all-pass filter, with a z-domain transfer function given by

$$\frac{V_{out}}{V_{in}} = \frac{z^{-1}C_1/C_2}{1 - z^{-1}(1 - C_1/C_2)} \tag{14}$$

which, when $C_3 = C_2$, reduces to $V_{out}/V_{in} = z^{-1}C_1/C_2$. Therefore, the input can be simply sampled and scaled by C_1/C_2, or a single-pole discrete-time filter can also be included if so desired. This circuit can easily be utilized with conventional switched-capacitor design and has even been successfully used for subsampling a 910 MHz input at a 78×10^6 samples/s rate (14).

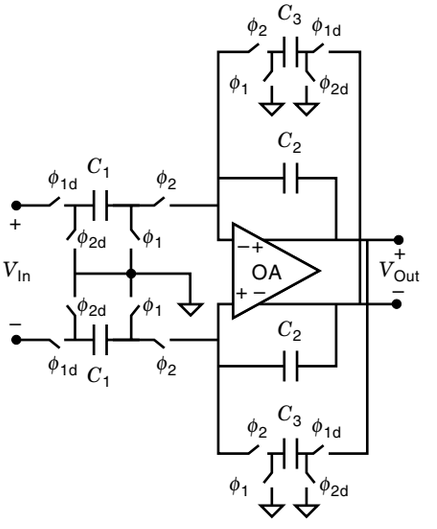

Figure 20. A first-order all-pass switched-capacitor filter used as a sample-and-hold circuit.

The sample-and-hold circuit shown in Fig. 20 differs from those shown in Figs. 18 and 19 in that the circuit's output is not reset during a clock phase. The output transitions from its previous value to its new value during ϕ_2 and then remains constant on ϕ_1, thus giving circuitry that follows longer time to process the signal. Another switched-capacitor sample-and-hold circuit with this same property is shown in fully differential form in Fig. 21 and utilizes both positive and negative feedback during ϕ_2. On ϕ_1, the input is sampled onto C_3, and the present output is sampled onto C_1. At ϕ_2, the C_3 capacitors are switched in parallel with the C_2 feedback capacitors, and the C_1 capacitors are switched in positive feedback around the amplifier. Using $C_1 = C_3$ results in $V_{out}/V_{in} = z^{-1}$, again without any reset phase necessary (15,16).

While the architectures discussed above all require at least two or three clock phases to function properly, in high-speed applications it is desirable to use as few clocks as possible, preferably just one. Figure 22 shows such a circuit that has achieved 150×10^6 samples/s operation in a 0.7 μm bipolar CMOS process (17). While ϕ_1 is high, the circuit operates as a continuous-time amplifier with transfer given by

$$\frac{V_{out}}{V_{in}} = -\frac{R_2}{R_1}\frac{1 + j\omega R_1 C_1}{1 + j\omega R_2 C_2} \tag{15}$$

which reduces to $V_{out}/V_{in} = -R_2/R_1$ when $R_1C_1 = R_2C_2$. When ϕ_1 falls, the input voltage is sampled and buffered to the output. As in the switched-capacitor designs, the sampling switch operates at the operational amplifier virtual ground voltage at all times and thus (to first order) does not introduce signal-dependent charge injection at sampling. A danger with this design is that the input may still feed through to the output even during hold mode via the capacitors and resistors. This is only a concern if the operational amplifier's output resistance is too low, but the effect can still be reduced by adding an extra switch to short out the differential input voltage during the hold mode, as shown by the dashed line in Fig. 22. While this makes a two-phase clocking scheme necessary, it may be required depending on the operational amplifier design and the hold-mode feed-through requirements.

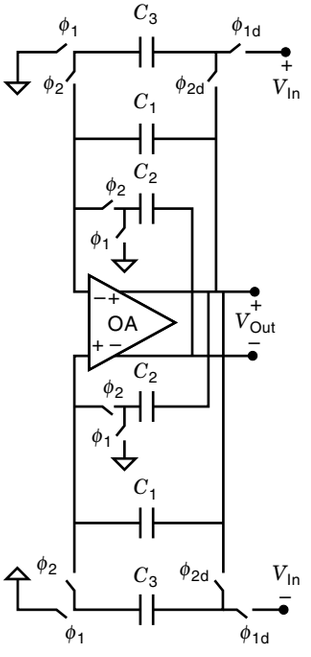

Figure 21. Another switched-capacitor circuit used as a sample-and-hold circuit, now using both positive and negative feedback around the operational amplifier.

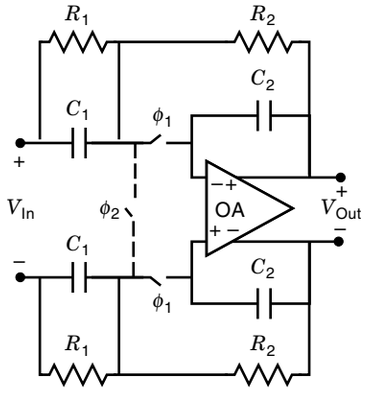

Figure 22. A sample-and-hold circuit design using closed-loop feedback during both sampling and hold modes and needing only a single clock phase (ϕ_2 is not necessary).

A similar design that also provides input buffering is shown in Fig. 23 and replaces the resistors in the previous design with a differential input transconductance amplifier with gain G (18,19). During acquisition mode, the system has a transfer function of

$$\frac{V_{out}}{V_{in}} = \frac{1}{1 + j\omega C/G} \tag{16}$$

that results in $V_{out} \approx V_{in}$ for frequencies well below $\omega_0 = G/C$. The sampling switch isolates the integrator from the transconductance amplifier during the hold mode, so the integrator holds its state until the switch is turned on again. As in the circuit of Fig. 22, the sampling switch always operates at an approximately constant virtual ground voltage (assuming the operational amplifier has sufficient gain and bandwidth for the application), so its charge injection and clock feed-through will be signal independent to first order. The closed-loop negative feedback reduces system nonlinearity caused by that of the open-loop transconductance amplifier and operational amplifier. An extra switch at the transcon-

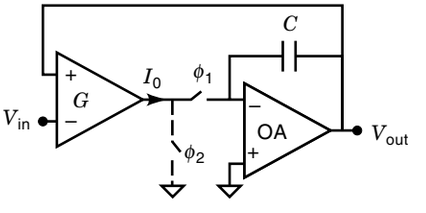

Figure 23. A closed-loop sample-and-hold circuit with both input and output buffering, using a differential input transconductance amplifier and a switched-capacitor integrator. The switch shown with the dashed line is recommended to prevent the transconductance amplifier output from saturating while ϕ_1 is off.

ductance amplifier output to ground is also recommended, in order to prevent that node from saturating high or low while the sample-and-hold circuit is in hold mode. Without this switch, the transconductance amplifier may experience a longer recovery time when the circuit returns to acquisition mode, thus lowering the circuit's maximum sampling rate.

In general, sample-and-hold circuit designs that utilize closed-loop feedback during sampling, such as that shown in Fig. 23, are not widely used. Input buffering can be achieved at very high linearity albeit moderate speed using a high-gain operational amplifier in unity-gain configuration, and open-loop sampling at the input of a switched-capacitor integrator can achieve linearity in excess of 100 dB when designed carefully. The use of an operational amplifier in closed-loop feedback in an integrator configuration is very common, with a variety of switching configurations possible, and is widely used to also provide output buffering. Closed-loop designs like that shown in Fig. 22 are often used not for the increased linearity of a feedback configuration but instead for the need of a phase with only a single clock, an advantage in high-speed design. Interestingly, a closed loop sample-and-hold topology as described in Ref. 20 can be even used to perform a single-ended to differential voltage conversion.

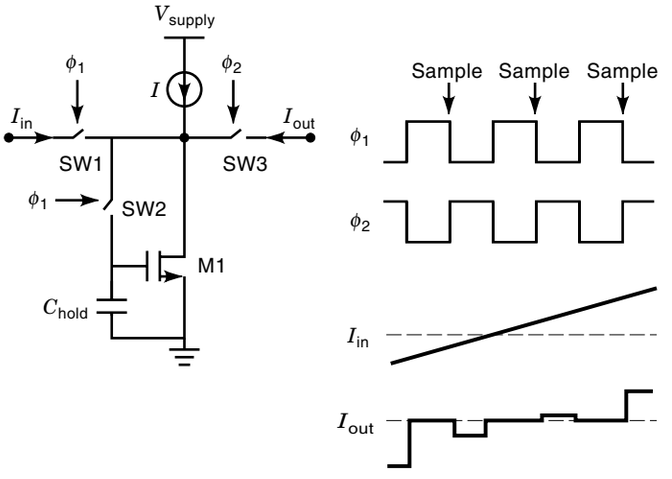

Figure 24. Current-mode sample-and-hold circuit (current copier). Capacitor C_{hold} is charged during ϕ_1 and holds its charge during ϕ_2 generating the output current I_{out} through transistor M1.

Current-Mode Sample-and-Hold Circuits

Current-mode sample-and-hold circuits provide an alternative to voltage-mode implementations described so far. Conversions between current and voltage usually influence the performance of the overall circuit, and selecting between the two types of sampling provides an additional degree of freedom during design. Currents could be stored using inductors similar to voltages being stored on capacitors. However, monolithic implementation of inductors is not practical in today's technologies and even discrete components do not offer a viable solution due to their nonzero series resistance, which leads to loss. Also, switching inductors without generating significant transients appears as extremely challenging.

The easiest and most common method for storing currents is to convert the current to a voltage or charge and store that on a holding capacitor. During the hold mode, the current can be restored through an inverse voltage-to-current conversion. Since the same transconductance element can be used for the current-to-voltage and the voltage-to-current conversions, the overall performance is not affected in a first-order approximation. This subsection will discuss some basic topologies for current-mode sample-and-hold circuits as well as their nonidealities.

Most current-mode circuits including sample-and-hold circuits are implemented using MOS technology. Very small gate currents (compared to significant base current in the bipolar case) combined with the existence of gate capacitance that could be used as a holding capacitor, easy switch, and logic function implementation are a few reasons behind this selection. For monolithic implementations, MOS technology is also less expensive. Figure 24 presents a basic current-mode sample-and-hold circuit called the alternatively current copier. Two phases of a clock control the circuit functionality. When ϕ_1 is high, switches SW1 and SW2 are closed and the input current I_{in} flows into transistor M1 together with the bias current I. During this period switch SW3 is opened and there is no current generated at the output $I_{out} = 0$. Since transistor M1 is configured as a MOS diode during this phase, the voltage on the holding capacitor C_{hold} is

$$V_{C_{hold}} = V_{GS,M1} = V_T + \sqrt{\frac{2(I + I_{in})}{\mu_n C_{ox} W/L}} \tag{17}$$

where V_{Th} and W/L are the threshold voltage and the size ratio of transistor M1, respectively, μ_n is the electrical mobility corresponding to n-channel devices, and C_{ox} is the oxide capacitance. When ϕ_1 goes low, the input current value at that instant is sampled as a voltage on C_{hold}, according to Eq. (17). When ϕ_2 goes high, switches SW1 and SW2 are opened and switch SW3 is closed. This is the hold phase, when the output current can be expressed as

$$I_{out} = \tfrac{1}{2}\mu_n C_{ox}(V_{GS,M1} - V_T)^2 - I = I_{in}|_{sampled} \tag{18}$$

since $V_{GS,M1} = V_{C_{hold}}$ as it was stored when ϕ_1 went low. It can be seen that the output current is a perfect copy of the sampled input current in the absence of nonidealities. However, charge injection from switches and finite output impedance of transistor M1 can seriously affect the performance. Also, it is important that ϕ_1 and ϕ_2 do not overlap since the sample stored on C_{hold} is corrupted if switches SW2 and SW3 are closed at the same time. Figure 25 shows a simple circuit that generates a two-phase nonoverlapping clock. The nonoverlap duration is given approximately by three logic gate delays.

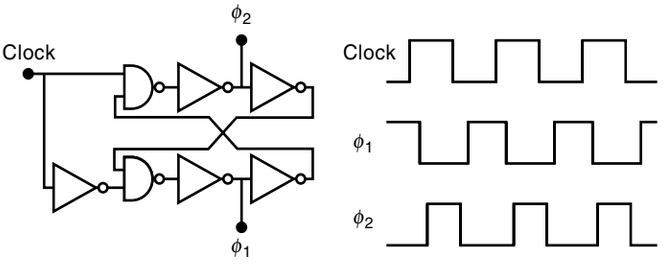

Figure 25. Circuit for generating nonoverlapping phases of a clock.

There are circuit techniques and topology variations that can be employed to attenuate the effect of nonidealities or to improve the speed. When the switch SW1 in Fig. 24 is turned off, a part of the charge stored in its channel is dumped into the holding capacitor. This charge is signal dependent and results in nonlinearity at the output. One method to alleviate this effect was described in the subsection concerning open-loop sample-and-hold circuits and relies on dummy-switch cancellation (21). If the clock edges are steep enough and switches are implemented using MOS transistors, half of the charge from switch SW1 appears on C_{hold} when this switch is turned off. Consequently, a dummy switch with half the width of switch SW1 can be turned on at the same time, which would attract the parasitic charge from C_{hold} generated by switch SW1. Figure 26 shows the dummy-switch cancellation technique. The size ratio of transistor M_{dummy} is half that of M1 since the parasitic channel charge of a MOS device is proportional to this ratio, as described in Eq. (4). Also, switches SW1 and SW3 were replaced with NMOS transistors M3 and M4 such that for the shown level of clock phases ϕ_1 and ϕ_2 they preserve the switch functionality from Fig. 24. For low and medium resolution (less than 8 bits) the dummy-switch cancellation provides a reliable solution and allows fast operation due to the small number of components, hence parasitics, involved. However, mismatches between the dummy and main switches, nonideal clock edges, and ultimately process-dependent channel charge present a serious challenge for higher-resolution designs. Employing differential topologies can alleviate the effect of the mismatches. Also, since the current error due to charge injection is inversely proportional to the holding capacitor, making C_{hold} larger improves accuracy at the expense of speed. This could be circumvented using Miller-type closed-loop structures that increase the effective capacitance value with a negligible penalty in speed performance (2).

Another way of reducing the effect of signal dependent parasitic charge injection is to perform a multiple-step cancellation. Figure 27 shows a two-step cancellation technique that uses multiple phases of a clock and was first described in Ref. 22. Phase ϕ_1 is split into ϕ_{11} and ϕ_{12}, which contribute to the

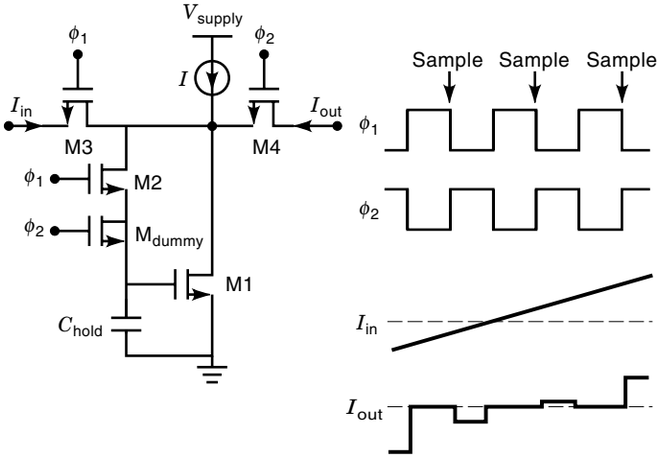

Figure 26. Charge-injection error cancellation using a dummy switch. The circuit is similar to the one described in Fig. 24, except switches were replaced by transistors M2, M3, M4, and a dummy (M_{dummy}) was added.

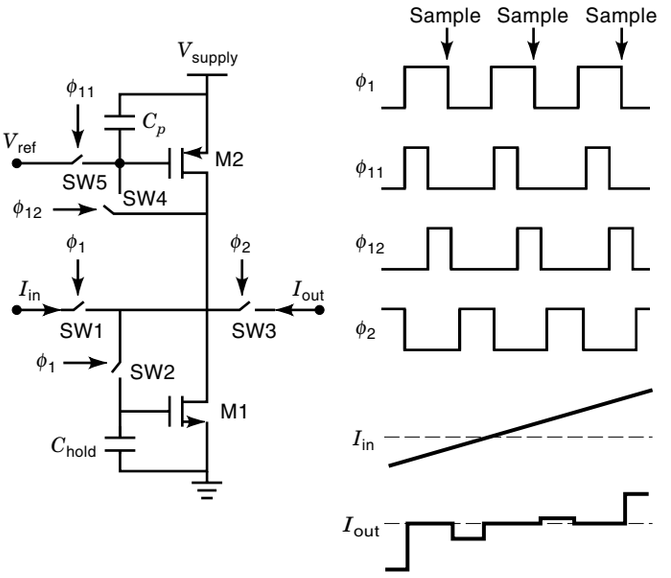

Figure 27. Two-step charge-injection error cancellation using multiple clock phases.

cancellation process. When ϕ_{11} is high, ϕ_1 is also high and the circuit behaves similarly to that in Fig. 24. Also, during this phase V_{ref} is connected to the gate of the PMOS transistor M2, which acts as a current source with the value I_{ref}. When ϕ_{11} goes low, the voltage on the holding capacitor C_{hold} corresponds to a current of $I_{ref} + I_{in} + I_{err}$ through transistor M1, where I_{err} is due to the charge injection from SW2. I_{err} exhibits a strong input signal dependency and if left uncanceled it would introduce nonlinearity in the output current. Phase ϕ_{12} goes high immediately following ϕ_{11}. During this, transistor M2 becomes diode connected and generates a current of $I_{ref} + I_{err}$. The corresponding voltage is stored on capacitor C_p (see Fig. 27). When ϕ_{12} goes low, ϕ_1 also goes low and transistors M1 and M2 remain with their gates connected only to C_{hold} and C_p, respectively. This is the hold phase: ϕ_2 is high and the output current is generated as a difference between the current through M1 and the current through M2. The current through M2 is $I_{ref} + I_{err} + I_{off}$, where I_{off} is due to the charge injection on C_p from switch SW4 being turned off and the end of ϕ_{12}. Since the current through transistor M1 remained as $I_{ref} + I_{in} + I_{err}$, the output current during the hold phase becomes $I_{in} + I_{off}$. The channel charge from switch SW4, which produces I_{off}, contains very little input signal dependency because the voltage across switch SW4 during ϕ_{12} is almost constant with respect to the input signal. This is valid only if the frequency of the clock that generates ϕ_1, ϕ_{11}, ϕ_{12}, and ϕ_2 is low enough compared to the signal bandwidth and the input current I_{in} varies very little during ϕ_{12}. As a result the output current is an accurate copy of the input plus an offset component.

A major performance degradation in current-mode sample-and-hold circuits is due to the finite input resistance of MOS transistors. This introduces nonlinearity since the current stored and generated by a MOS device acquires a component that is a function of the drain-to-source voltage. This voltage is more often signal dependent unless some precautions are employed. Cascode structures and differential topologies can be used to reduce this nonideality. Also, Ref. 23 shows the

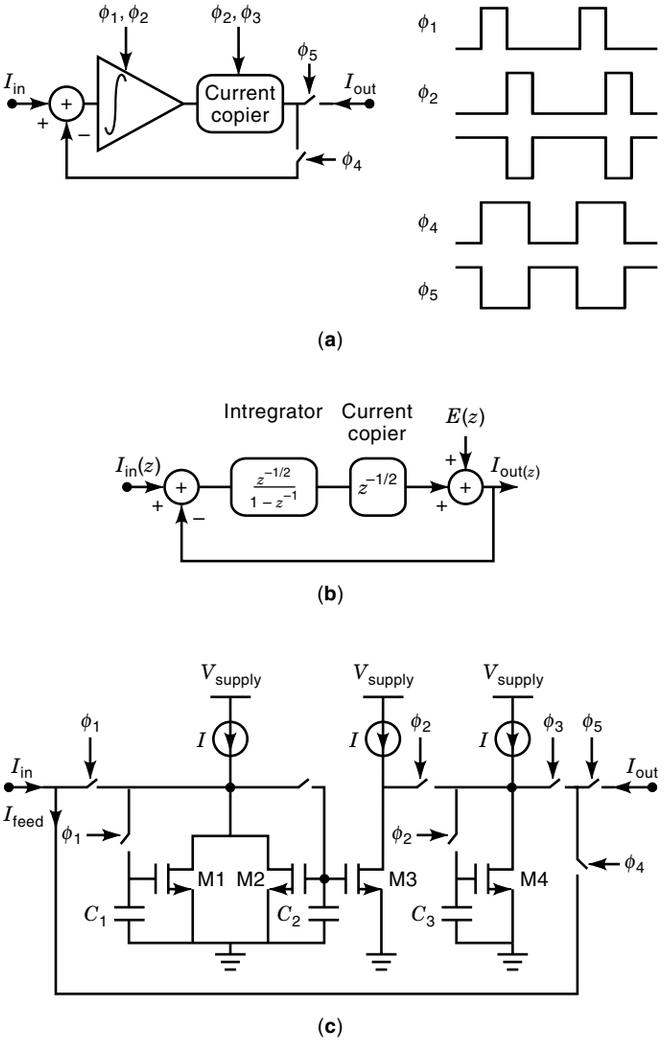

Figure 28. Current-mode sample-and-hold circuit using negative feedback and oversampling. (a) Conceptual diagram; (b) z-domain block diagram; (c) transistor level implementation.

usage of negative feedback to increase the output resistance of a MOS device.

Negative feedback and oversampling of the input signal can be used to reduce the effect of all types of errors described so far and implement a highly accurate current-mode sample-and-hold circuit at the expense of more hardware. Figure 28(a) presents a conceptual diagram for such a circuit (24). The difference between the input and output currents is fed into a current-mode integrator clocked by ϕ_1 and ϕ_2. Its output is copied to the overall circuit output if ϕ_5 is high or fed back to create the aforementioned difference during ϕ_4. Clearly, the output is an accurate copy of the sampled input since the integrator and the closed loop would force their difference to zero. Before describing the functionality in more detail, it is useful to observe the circuit from a z-domain perspective as in Fig. 28(b). The integrator and the current copier are assumed to be ideal and their z-domain transfer functions are $z^{-1/2}/(1 - z^{-1})$ and $z^{-1/2}$, respectively. The term $z^{-1/2}$ in the integrator transfer function signifies that its output is available during ϕ_2 and the input is sampled at the end of ϕ_1. Any error in the current copier can be mapped as $E(z)$ and added into the signal path as shown in Fig. 28. The transfer functions from the input $I_{in}(z)$ and $E(z)$ to the output $I_{out}(z)$ can be expressed as

$$H(z) = \frac{I_{out}(z)}{I_{in}(z)} = z^{-1}, \qquad H_\epsilon = \frac{I_{out}(z)}{E(z)} = 1 - z^{-1} \tag{19}$$

It can be seen that the input transfer function $H(z)$ is just a delay as in the case of any sample-and-hold circuit. On the other side, the error transfer function H_ϵ has a zero at dc (when $z = 1$) similar to the noise-shaping performed by a first-order Δ–Σ data converter. As a result, H_ϵ is a high-pass characteristic and the amount of error at the output is proportional to the ratio of the input frequency f_{input} to the frequency of the clock f_{clock}. The errors associated with the integrator are also attenuated by the aforementioned oversampling ratio (OSR = $f_{clock}/2f_{input}$) as discussed in Ref. 24. By increasing the number of integrators and feedback loops, the analogy to Δ–Σ modulation can be continued and the accuracy of the output current increases accordingly. The speed can be also improved using parallel architectures based on the same circuit from Fig. 28(a) clocked on alternate phases.

Figure 28(c) shows a MOS transistor implementation of the diagram from Fig. 28(a). The difference between the input current I_{in} and the feedback current I_{feed} is applied to the integrator during ϕ_1. The functionality of the current-mode integrator can be easily explained assuming a constant input current I_{ct}. Assuming that initially the same current $I/2$ flows through both transistors M1 and M2, when ϕ_1 goes high transistor M1 will conduct $I/2 + I_{ct}$. When ϕ_1 goes low and ϕ_2 goes high transistor M2 will conduct $I/2 - I_{ct}$. When ϕ_1 goes high again, M1 will conduct $I/2 + 2I_{ct}$ and correspondingly when ϕ_2 goes high transistor M2 will conduct $I/2 - 2I_{ct}$. Clearly, the left-hand branch containing transistor M1 will conduct more and more current, while the right-hand branch including transistor M2 will conduct less and less. Transistor M3 mirrors out the current resulting from the integration process. Then, transistor M4 is used to create a current copier as in Fig. 24, which generates alternatively the output and the feedback currents.

Another circuit that uses feedback to provide an accurate sample of the input current is shown in Fig. 29 (25). This is similar to Fig. 23 except that instead of a voltage at the input of the transconductance, a current at its output is sampled.

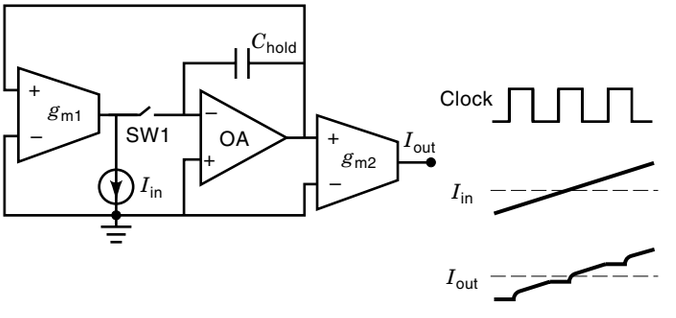

Figure 29. Closed-loop current-mode sample-and-hold circuit. This circuit resembles the one in Fig. 23, except that a current is stored as opposed to a voltage.

When the clock is high the input current is stored as a voltage on C_{hold}. Since the negative input to the operational amplifier OA is preserved as a virtual ground, the feedback loop containing g_{m1}, switch SW1, and the operational amplifier OA forces the voltage on the holding capacitor to $V_{C_{\text{hold}}} = I_{\text{in}}/g_{\text{m1}}$ where g_{m1} is the transconductance value. During the hold period when the clock signal is low, the output is maintained at

$$I_{\text{out}} = \frac{g_{\text{m2}}}{g_{\text{m1}}} I_{\text{in}}|_{\text{sampled}} \tag{20}$$

It is worth mentioning that g_{m2} and g_{m1} are physically different elements and their mismatches as well as nonlinearity impact the performance of this sample-and-hold circuit. Also, the speed of the circuit depends largely on the operational amplifier settling time.

Real-life implementations of the current-mode sample-and-hold circuits described so far include usually more elaborate structures that are targeted at cancelling parasitic effects. Some of them employ differential topologies or improvements related to the finite output impedance of MOS transistors. Other structures employ more sophisticated clocking schemes or feedback to cancel charge-injection errors. Furthermore, closed-loop architectures must deal with stability issues that come into play as a trade-off regarding the accuracy (in the case of discrete-time feedback as in Fig. 28) or the speed (for the continuous-time feedback as in Fig. 29) of the circuit. Although there are advantages in using current-mode sample-and-hold circuits particularly regarding operating speed, their widespread usage was hampered by the need of voltage-to-current and current-to-voltage conversions as well as implementation difficulties when higher accuracy is required.

CONCLUSIONS

Sample-and-hold circuits are widely used in a broad variety of circuits and systems, from sensor-conditioning applications to communications designs to data-conversion systems. In many cases the whole system design starts with specifying the sample-and-hold operation, and its implementation can even dictate a specific overall architecture. While the concept of a sample-and-hold circuit seems quite simple in its idealized form, the previous sections have shown how complex and detailed the designs and requirements can become. The various performance metrics that have been established for evaluating sample-and-hold operation are useful in setting specifications for a design to be used in particular applications. These specifications then allow the designer to survey the many techniques and architectures that exist for sample-and-hold circuits and focus on the one that best meets these requirements and is most efficient for implementation. The broad variety of design topologies available is evidence itself of their differing advantages and disadvantages, all of which must be considered when determining the suitability of a design for the given application.

The designs discussed here were chosen as a representative subset of the many sample-and-hold architectures that exist, and myriad variations and modifications on these can be found. The techniques used in the designs depend heavily on the integrated-circuit technology, with series switches and switched-capacitor-based designs most prevalent in MOS technologies, while current steering and diode bridges are common in bipolar designs. Current-based designs, as discussed earlier, are attractive for applications involving current-output sensors and current-mode signal processing, and have given rise to several innovative design concepts. Higher-speed and finer-line integrated-circuit processes will continue to push existing design techniques to higher and higher performance levels, and innovative conceptual and architectural breakthroughs will provide an extra boost to exceed the performance barriers that rise as time goes on.

BIBLIOGRAPHY

1. W. Kester and J. Bryant, Sample and Hold Circuits in W. Kester (ed.), *Analog Devices Linear Design Seminar,* Norwood, MA: Analog Devices, 1995, pp. 8.44–8.67.
2. B. Razavi, *Principles of Data Conversion System Design,* Piscataway, NJ: IEEE Press, 1995.
3. A. Sanchez, Understanding sample-hold modules, in *Analog Devices Analog Dialogue,* Norwood, MA: Analog Devices, 1971, pp. 51–54.
4. P. E. Allen and D. R. Holberg, *CMOS Analog Circuit Design,* Orlando, FL: Saunders HBJ, 1987.
5. T. L. Brooks et al., A 16b ΣΔ pipeline ADC with 2.5 MHz output data-rate, *ISSCC Dig. Tech. Papers,* San Francisco, 1997, pp. 208–209.
6. G. Wegmann, E. A. Vittoz, and F. Rahali, Charge injection in analog MOS switches, *IEEE J. Solid-State Circuits,* **SC-22** (6): 1091–1097, 1987.
7. T. V. Burmas et al., A second-order double-sampled delta-sigma modulator using additive-error switching, *IEEE J. Solid-State Circuits,* **31** (3): 284–293, 1996.
8. P. Vorenkamp and J. P. M. Verdaasdonk, Fully bipolar, 120-Msample/s 10-b track-and-hold circuit, *IEEE J . Solid-State Circuits,* **27** (7): 988–992, 1992.
9. B. Razavi, Design of a 100-MHz 10-mW 3-V sample-and-hold amplifier in digital bipolar technology, *IEEE J. Solid-State Circuits,* **30** (7): 724–730, 1995.
10. B. Razavi, A 200-MHz 15-mW BiCMOS sample-and-hold amplifier with 3 V supply, *IEEE J. Solid-State Circuits,* **30** (12): 1326–1332, 1995.
11. M. H. Wakyama et al., A 1.2-μm BiCMOS sample-and-hold circuit with a constant-impedance, slew-enhanced sampling gate, *IEEE J. Solid-State Circuits,* **27** (12): 1697–1708, 1992.
12. T. Wakimoto and Y. Akazawa, Circuits to reduce distortion in the diode-bridge track-and-hold, *IEEE J. Solid-State Circuits,* **28** (3): 384–387, 1993.
13. P. E. Allen and E. Sanchez-Sinencio, *Switched-Capacitor Circuits,* New York: Van Nostrand Reinhold, 1984.
14. D. H. Shen et al., A 900 MHz RF front-end with integrated discrete-time filtering, *IEEE J. Solid-State Circuits,* **31** (12): 1945–1954, 1996.
15. K. Y. Kim, N. Kusayanagi, and A. A. Abidi, A 10-b, 100MS/s CMOS A/D converter, *IEEE J. Solid-State Circuits,* **32** (3): 302–311, 1997.
16. G. Nicollini, P. Confalonieri, and D. Senderowicz, A fully differential sample-and-hold circuit for high-speed Applications, *IEEE J. Solid-State Circuits,* **24** (5): 1461–1464, 1989.
17. L. Schillaci, A. Baschirotto, and R. Castello, A 3-V 5.4-mW BiCMOS track & hold circuit with sampling frequency up to 150 MHz, *IEEE J. Solid-State Circuits,* **32** (7): 926–932, 1997.
18. U. Gatti, F. Maloberti, and G. Palmisano, An accurate CMOS sample-and-hold circuit, *IEEE J. Solid-State Circuits,* **27** (1): 120–122, 1992.

19. K. R. Stafford, A complete monolithic sample/hold amplifier, *IEEE J. Solid-State Circuits,* **SC-9** (6): 381–387, 1974.
20. W.-C. Song et al., A 10-b 20-Msample/s low-power CMOS ADC, *IEEE J. Solid-State Circuits,* **30** (5): 514–521, 1995.
21. T. S. Fiez and D. J. Allstot, Switched-current circuit design issues, *IEEE J. Solid-State Circuits,* **26** (3): 192–202, 1991.
22. J. B. Hughes and K. W. Moulding, S^2I: A switched-current technique for high performance, *Electron. Lett.,* **29** (16): 1400–1401, 1993.
23. X. Hu and K. W. Martin, A switched-current sample-and-hold circuit, *IEEE J. Solid-State Circuits,* **32** (6): 898–904, 1997.
24. I. Mehr and T. L. Sculley, A 16-bit current sample/hold circuit using a digital CMOS process, *Proc. IEEE Int. Symp. Circuits Syst.,* 1994, pp. 5.417–5.420.
25. D. H. Robertson, P. Real, and C. Mangelsdorf, A wideband 10-bit, 20 Msps pipelined ADC using current-mode signals, *ISSCC Dig. Tech. Papers,* San Francisco, 1990, pp. 160–161.

IURI MEHR
Analog Devices, Inc.
TERRY SCULLEY
ESS Technology, Inc.
NADI ITANI
Cirrus Logic, Inc.

SAMPLED-DATA CONTROL. See DIGITAL CONTROL.
SAMPLED-DATA SYSTEMS. See SWITCHED CAPACITOR NETWORKS.
SAMPLING WATTMETERS. See WATTMETERS.