

Nonlinearity Correction Techniques for High Speed, High Resolution A/D Conversion

Terry L. Sculley, *Member, IEEE* and Martin A. Brooke, *Member, IEEE*

Abstract—In general, low resolution circuitry can be designed with higher speed performance than high resolution circuitry. Most high resolution A/D converter architectures require a high resolution component, such as a S/H or integrator, in their design and are often limited by the speed of this component. This paper discusses several techniques for achieving high resolution A/D conversion by calibrating nonlinear circuitry with highly accurate reference circuitry without conversion speed being limited by the calibration circuitry. The analysis of a sample converter with various nonidealities is presented, along with supporting simulation results for a similar architecture.

I. INTRODUCTION

OVER THE YEARS, a wide variety of A/D converter architectures have been used to exploit the strengths of a particular technology and to address the requirements of specific applications. As a converter topology was pushed to the limits of its achievable performance, new techniques were often developed to overcome those limitations and extend performance to even higher levels. In the case of high resolution converters, monolithic successive-approximation architectures were limited in resolution due to mismatch between devices on chip, but the application of self-calibration techniques extended their resolution [1], [2]. Similarly, the use of noise-shaping topologies in monolithic converter designs pushed resolution to even higher levels by overcoming both noise and mismatch limitations [3]–[5].

Unfortunately, the speed performance of virtually all high resolution A/D architectures is still limited by highly accurate but slower components, such as a highly linear sample-and-hold (S/H) or integrator, embedded in the converter's critical timing path. This is especially true in the case of current-mode converters, where highly linear current S/H's are difficult to implement. Often a higher speed circuit can be designed at lower resolution, but most A/D architectures cannot tolerate this reduced resolution performance. Thus, speed improvements in high resolution A/D's have often come through the design of faster components that still maintain the needed high accuracy [6].

Another approach to improving conversion speed is to modify the converter architecture to remove the highly accurate

components from the critical timing path. Then the A/D's speed is determined by lower resolution circuitry, which can be designed with higher speed much more easily. Since the speed of the highly accurate circuitry is no longer as critical, the design of these components can focus instead on achieving higher accuracy and possibly increasing A/D resolution even further. In addition, the ability to use a faster but possibly nonlinear S/H could increase the achievable analog input bandwidth of the converter at the same time [7]–[9].

This paper focuses on techniques for overcoming nonlinearity in the various components (S/H, residue amplifiers, D/A's) in the critical timing path of an A/D converter. Section II discusses performing a full mapping over the S/H output, while Section III shows the use of piecewise-linear approximations. In Section IV, the operation of a sample A/D converter is explained, and analytical results are presented, with the effects of various nonidealities included. Simulation results for a similar converter architecture are shown in Section V, some implementation considerations in Section VI, and conclusions in Section VII.

II. A/D MAPPING

The most obvious approach to overcoming circuit nonlinearities would be to use modified self-calibration techniques to correct or linearize the transfer characteristic of a high speed nonlinear A/D converter. A simple successive approximation A/D converter architecture that embodies this idea is shown in Fig. 1. The RAM stores corrections to the normal successive approximation codes to correct for nonlinearity, and the main D/A translates those codes to a corrected analog output signal for comparison. Although a highly accurate D/A converter is still required for calibration, it is taken out of the critical timing path of the converter, enabling the D/A to operate at low speed without limiting the overall A/D conversion speed. The circuitry remaining in the main loop now has reduced precision requirements, making high speed operation easier to obtain. The primary remaining restriction on the S/H is that it must have a repeatable, monotonic transfer function, but high linearity is not required. The main D/A must have a large number of input bits ($>N$), but it does not require the corresponding high level of linearity in its transfer characteristic. Its main requirement is that it have sufficiently fine adjustability over the S/H output range, which is easily accomplished by using redundancy in its design without requiring high circuit precision [10].

The memory requirements for this converter architecture are approximately $N2^N$ bits, with some extra memory necessary

Manuscript received September 9, 1993; revised May 5, 1994. The work of T. L. Sculley was supported by a Semiconductor Research Corporation Graduate Fellowship and a Georgia Tech President's Fellowship. This paper was recommended by Associate Editor L. R. Carley.

T. L. Sculley is with the School of Electrical Engineering and Computer Science, Washington State University, Pullman, WA 99164 USA.

M. A. Brooke is with the Microelectronics Research Center, Georgia Institute of Technology, Atlanta, GA 30332 USA.

IEEE Log Number 9407656.

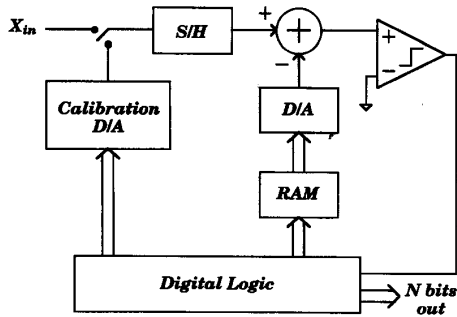


Fig. 1. Successive-approximation A/D converter using nonlinearity correction.

to ensure adequate adjustability in the main D/A. Therefore, a 16 b A/D converter would require over 1 Mb of RAM. This large memory requirement makes this conversion technique impractical for a high resolution monolithic IC implementation. While this technique could also be applied in a parallel or combination serial/parallel converter architecture, the digital memory needed remains the same.

By modifying the converter architecture of Fig. 1 to utilize the existing linearity of the S/H and main D/A, the amount of memory needed can be reduced. This is implemented by digitally summing the ideal main D/A input code with a stored correction value to get the final main D/A input code. Although the memory required drops as the S/H and main D/A linearities increase, one or more correction bits are still required for each of the A/D's 2^N output codes. Thus, this technique remains impractical for a high resolution monolithic implementation.

III. PIECEWISE-LINEAR APPROXIMATION

In the previous section, the techniques discussed overcome A/D nonlinearity using a point-by-point mapping from digital memory to every decision point between quantization levels. However, there is a useful characteristic of most S/H's and residue amplifiers that can be exploited to reduce memory size even further. The transfer curves of most of these components are smooth, with only slowly varying nonlinearities. Performing a piecewise-linear (PWL) approximation to the transfer curve, instead of a full mapping, can greatly reduce the number of bits that must be stored. In an A/D converter, this involves matching the linear output characteristic of a D/A converter to the amplifier output curve over a small range of the output. The length of the PWL segments is determined by the amplifier nonlinearity and by the D/A output nonlinearity. These constraints are illustrated in Fig. 2, in which a plot of S/H output versus analog input X_{in} is superimposed on a plot of D/A output versus digital input. The transition point between digital codes occurs when the D/A output equals the S/H or amplifier output, so to achieve N bit linearity for the converter the transition points must not deviate from the ideal by more than $1/2$ LSB. If the amplifier output curves too much or the D/A output is too nonlinear, the deviation can exceed $1/2$ LSB, meaning that the PWL segments must be shortened until the deviation is acceptable.

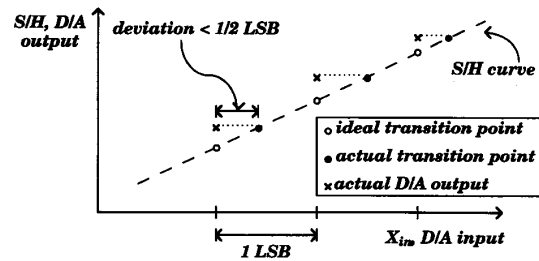


Fig. 2. Relation of D/A output to switching point of converter.

In order to match a D/A output characteristic to a S/H or residue amplifier output segment, the D/A output must be adjusted to start at the appropriate value (i.e., have the correct offset) and to have the proper slope. The offset is easily adjusted either by adding the output of an offset D/A to the main D/A output or by adding a digital offset code to the main D/A input code. Digitally adding an offset code constrains the main D/A output characteristic to be linear or have smooth nonlinearities over the entire S/H or amplifier output range, while adding an analog offset only constrains the main D/A to be linear over one short segment. For this reason, adding the analog offset is the preferred method.

Adjustment of the slope of the D/A output characteristic requires multiplication of the D/A output by a correction factor corresponding to the appropriate PWL segment. This multiplication can be performed in either an analog or a digital fashion. The analog solution adjusts the reference of a multiplying D/A (MD/A) converter in order to scale the output signal independent of the digital input code. Digital slope correction can be performed either before or after conversion of the lower bits. In the former case, the lower bits of the D/A input code are multiplied by a digital correction factor before being applied to the D/A, while in the latter case the lower bits of the A/D output are digitally scaled after conversion.

Block diagram implementations of these approaches to slope correction are shown in Fig. 3. The analog approach of Fig. 3(a) uses an extra D/A to adjust the MD/A reference, enabling the use of digital memory and calibration for robustness. The digital correction depicted in Fig. 3(b) was proposed in [11] for application to a successive approximation A/D converter. In that case, the digital multiplier could be implemented serially without significant speed penalty due to the nature of the converter's binary search operation. In more parallel designs, however, this technique becomes impractical in its present form. For those cases, the digital approach of Fig. 3(c) is more attractive, although a digital multiplier with high throughput rate is still necessary to avoid limiting conversion speed.

A brute force method of PWL correction would be to use a digital postprocessor after a nonlinear A/D converter. However, in order to ensure the desired linearity, more bits must be converted by the original nonlinear converter than are needed. In addition, the original converter must also have a smoothly varying transfer characteristic between PWL segments, with abrupt nonlinearities constrained only to the endpoints of the segments. These characteristics, along with

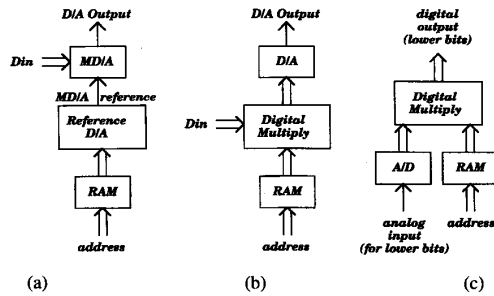


Fig. 3. (a) Analog slope adjustment of D/A output. (b) Digital slope adjustment of D/A output. (c) Digital slope adjustment of A/D output.

the still present need for a high speed digital multiplier, make this method unattractive in many situations.

Although these calibration techniques manage to overcome static nonlinearities in various A/D components, they do not address dynamic nonlinearities due to the speed of the input applied and to incomplete settling. However, now that high static linearity is no longer necessary for these components, their design can instead focus on increasing their speed performance. Then dynamic nonlinearities will become significant at higher speeds than before, thus improving converter speed performance.

Even by using high speed, low precision components in many parts of a converter, a 16 b converter with the same speed as the fastest 8 b converter is still not possible. The low precision circuitry in the design must still settle to the same nonideal performance to within 16 b levels, requiring longer settling times than if in an 8 b converter. Because of this, higher speed will not be possible for all low precision components. For example, a D/A that may now have large nonlinearity may be easier to design than a highly linear version, but the time required for precise settling will be comparable. However, a significant speed improvement can be realized in components such as the S/H or residue amplifier. In these designs, high linearity specifications generally require the use of feedback circuitry, which often must be compensated to ensure stability. When these linearity specifications are reduced enough, an open loop design can be used instead, yielding higher speed operation and settling.

An example A/D converter architecture using analog offset and slope adjustment is shown in Fig. 4. Although a parallel or pipelined architecture would have higher conversion speed potential, a successive approximation architecture was chosen for its simplicity. The potential for an increased analog input bandwidth would be the same for all the converters, due to their ability to use a higher speed, nonlinear S/H. The memory requirements of this architecture depend on both the number of bits to be converted and the accuracy or linearity of the D/A's and S/H used. For a 16 b converter implementation using components with approximately 8 b linearity (except the calibration D/A), the required memory drops to 3.6 kbytes. This value is now much more practical and can be reduced further by summing the ideal D/A input code with a RAM correction at the offset D/A as discussed in Section II.

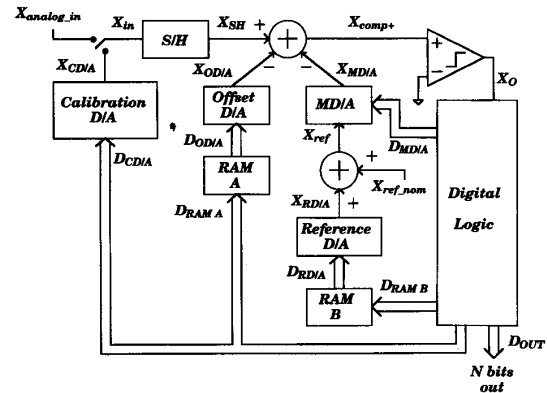


Fig. 4. Successive-approximation style A/D converter using PWL approximation.

Each of the converter architectures of Figs. 1 and 4 will be limited in resolution by the calibration D/A. This D/A must have resolution slightly greater than that desired of the A/D, with the exact requirement dependent on the converter's error budget. A 16 b A/D converter may likely require an 18 or 19 b calibration D/A to ensure linearity. However, because this high precision D/A has been removed from the critical timing path of the converter, this component can be designed for very high resolution at low speed without limiting the conversion speed [12]–[20].

In order to calibrate these converters, the input signal must be disconnected and the calibration signal applied. Calibration can be done either all at once, in a burst mode, or in many short periods, in an interleaved fashion. Although a burst calibration may take a significant amount of time to complete, after it is done the converter can run at full speed until another calibration is required. An interleaved calibration, however, allows the converter to maintain a constant but lower sampling rate, typically half of full speed, without interruption. In addition, during the time the calibration D/A is changing states and settling, the converter can be converting and just not using the cycles reserved for calibration. The analog input bandwidth stays the same for both burst and interleaved calibration, but the maximum sampling rate is lower for the interleaved case. One added note about the interleaved calibration case is that excessive perturbation of the calibration D/A output must be avoided during sampling. Otherwise, a much longer sampling period will have to be used to allow the D/A to settle back again. If this is a problem, modified sampling techniques [21] can be used to reduce the level of perturbation and thus the settling time required.

IV. OPERATION

The A/D converter of Fig. 4 determines the top $N - L$ bits using the offset D/A and the lower L bits using the MD/A. This effectively divides the S/H transfer curve into 2^{N-L} PWL segments, each of length 2^L LSB's. The endpoints of the segments correspond to the ideal transition points of the top $N - L$ bits.

During conversion, a binary search is performed using RAM A and the offset D/A to convert the top $N-L$ bits, as in Fig. 1. These top $N-L$ bits are then used, along with RAM B and the reference D/A, to adjust the MD/A slope. Next, another binary search is performed using the MD/A inputs to determine the lower L bits.

Calibration is performed by applying accurate calibration inputs to the converter corresponding to the endpoints of the PWL segments. When each calibration input is applied using $D_{CD/A} = k2^{N-L}$, $k \in [0, 2^L - 1]$, $D_{MD/A}$ is set to zero and a binary search is performed using the offset D/A and the comparator. The value of $D_{OD/A}$ resulting from the binary search is then stored in RAM A. Next, the calibration input is changed to $D_{CD/A} = (k+1)2^{N-L}$, and, with $D_{OD/A}$ still left at its previous value, $D_{MD/A}$ is set to 2^L and a binary search is performed using the reference D/A, the MD/A, and the comparator. The value of $D_{RD/A}$ resulting from the binary search is then stored in RAM B.

A mathematical analysis of the operation of Fig. 4 was performed, with nonidealities associated with the S/H and the offset, reference, and multiplying D/A's, along with comparator offset, included in the analysis. Errors in the calibration D/A were assumed negligible, along with comparator finite gain effects. The various equations used to model the components in Fig. 4 are given in Table I. The MD/A is assumed to have one extra LSB that can be added to the output, enabling the digital input to range over $[0, 2^L]$ instead of just $[0, 2^L - 1]$. The S/H transfer curve is modeled with an offset, a linear gain, and the function $f_{SH}(X_{in})$, which includes all higher order nonlinearities. An error budget was calculated for the A/D architecture by constraining the A/D input value at which the converter switches from one digital code to the next higher code to be within 0.5 LSB of its ideal value. The analysis is detailed in the Appendix. The resulting error budget equation is given by

$$\begin{aligned} & \frac{2^N}{A_{SH}(X_{in,max} - X_{in,min})} \\ & \times \left\{ \max \left| f_{SH}(X_{in} : n2^L) + \frac{m}{2^L} [f_{SH}(X_{in} : (n+1)2^L] \right. \right. \\ & \quad \left. \left. - f_{SH}(X_{in} : n2^L) \right| - f_{SH}(X_{in} : n2^L + m) \right| \\ & \quad + \max[|Q_{OD/A}(D_{RAM A})|, |Q_{RD/A}(D_{RAM B})|] \\ & \quad + \max[|E_{MD/A}(D_{MD/A}, D_{RAM B})|] \left. \right\} < \frac{1}{2}, \\ & \forall n \in [0, 2^{N-L} - 1] \text{ for every } m \in [0, 2^L - 1] \end{aligned} \quad (1)$$

where $X_{in:i}$ refers to the ideal input value at which the converter output code should switch from $D_{out} = i - 1$ to $D_{out} = i$, and $Q_{OD/A}$ and $Q_{RD/A}$ refer to the quantization error at the output of the offset and reference D/A's, respectively, due to their finite resolution.

The first maximum term in (1) refers to the difference in the S/H output curve and the PWL approximation to it, using endpoints at $X_{in:n2^L}$ and $X_{in:(n+1)2^L}$. The next maximum term involves the quantization errors of the offset and reference

TABLE I
MODEL EQUATIONS FOR THE A/D CONVERTER OF FIG. 4

$X_{SH}(X_{in}) = O_{SH} + A_{SH}X_{in} + f_{SH}(X_{in})$
$X_{in:k}$ is shorthand for $X_{in} = X_{in,min} + (X_{in,max} - X_{in,min})\frac{k}{2^N}$
$X_{MD/A}(D_{MD/A}, D_{RAM B}) = \frac{D_{MD/A}}{2^L} X_{ref}(D_{RAM B}) + O_{MD/A} + E_{MD/A}(D_{MD/A}, D_{RAM B})$ where $E_{MD/A}(0, D_{RAM B}) = E_{MD/A}(2^L, D_{RAM B}) = 0$ by definition
$X_{CD/A}(D_{CD/A}) = X_{in,min} + (X_{in,max} - X_{in,min})\frac{D_{CD/A}}{2^N}$
$X_{OD/A}(D_{RAM A}=k)$ refers to the offset D/A output when its input is specified by the value stored in RAM A at location k.
$X_{RD/A}(D_{RAM B}=k)$ refers to the reference D/A output when its input is specified by the value stored in RAM B at location k.
$X_{ref}(D_{RAM B}=k) = X_{ref,nom} + X_{RD/A}(D_{RAM B}=k)$
$X_O = \begin{cases} 1, & X_{comp+} \geq O_{comp} \\ 0, & X_{comp+} < O_{comp} \end{cases}$

D/A's, and the final term is the straight line error of the MD/A output in LSB's.

This error budget can now be applied to the case of a 16 b converter. If a 6 b MD/A with 8 b linearity is used, its straight line error ($E_{MD/A}$) will fall within $\pm 1/8$ LSB. The adjustability of the offset and reference D/A's sets their quantization errors. If this maximum error is chosen to be $1/8$ LSB, then the offset D/A would require $\log_2(\text{range/adjustability})$ bits if no mismatch is present. For $N = 16$, this becomes $\log_2(2^{16}\text{LSBs}/(1/8\text{LSB})) = 19$ b. However, due to nonidealities, extra bits are needed for redundancy. With only 8 b linearity available, 1 to 2 b of redundancy must be included in the analysis to ensure the $1/8$ LSB adjustability, yielding a 20 to 21 b offset D/A. Similarly, the choice of an 8 b reference D/A provides $1/8$ LSB adjustability over a ± 16 LSB region when translated to the MD/A full scale output. So to satisfy the error budget, the S/H linearity error over each PWL segment now must not exceed $\pm 1/4$ LSB. It should be noted that the 20 b D/A will not have 20 b linearity; instead, its overall linearity will be approximately 8 b. The 20 input bits ensure the desired $1/8$ LSB adjustability, which means that any analog value within the D/A output range will always fall within $1/8$ LSB of some D/A output value less than or equal to it.

The total amount of RAM required by the converter depends on the linearity and resolution desired, as well as on the level of matching errors in the circuitry. Assuming a k bit D/A can be built with uncalibrated error within $\pm 1/2$ LSB, then the offset D/A for an N bit converter would require approximately $(N + 2 + N/k)$ input bits, including the required redundancy. Since a correction code is needed for each PWL segment endpoint, approximately $2^{N-L}(L + 2 + N/k)$ bits must be stored for the offset corrections, if 2^{N-L} segments are used. The number of bits required for the reference D/A depends on how much the slope of the S/H varies over its range, increasing

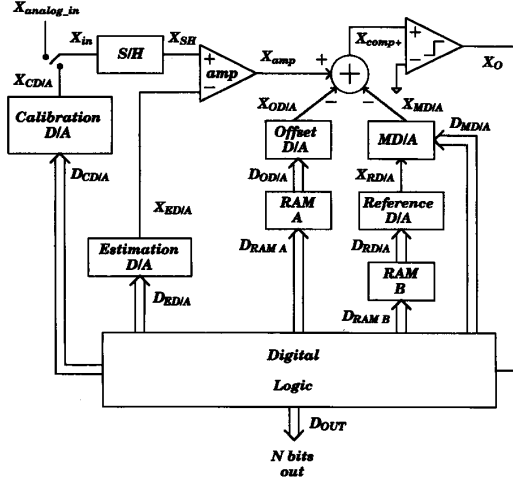


Fig. 5. Modified A/D converter architecture using PWL approximation.

with increasing S/H nonlinearity. In addition, the number of PWL segments varies depending on S/H nonlinearity. A highly nonlinear S/H will require many short PWL segments to fit it with low error, while the number of segments is reduced if higher error is acceptable. This demonstrates the trade-off possible between RAM size and final converter linearity, which is measured as differential nonlinearity (error in the difference between adjacent switching points) and integral nonlinearity (error of the switching points to a straight-line approximation). Since one MD/A reference adjustment is needed per segment, approximately $2^{N-L}(J)$ bits are required here, with J the number of bits of the reference D/A.

The above example demonstrates the capability of the proposed converter architecture to achieve high resolution using lower resolution components and one highly accurate but slow calibration component. Because the time for conversion depends on the speed of the low resolution components, not the slower high resolution calibration D/A, high conversion speed is possible. In addition, the size of the RAM required is inversely related to the accuracy and linearity attainable in the S/H and D/A's (except the calibration D/A), so that the RAM size drops as the accuracy increases.

Fig. 5 shows a modified version of the architecture of Fig. 4, with an estimation D/A and an extra amplifier included after the S/H. The use of the estimation D/A helps reduce the range over which the offset D/A must adjust, thus reducing the number of bits needed, and the amplifier helps reduce the noise sensitivity of the circuitry past it. For each of the top M bits ($M \leq N - L$), the top M bits of the digital code used to address RAM A are applied to the estimation D/A during conversion. Then the offset D/A is adjusted to set the switching point for the bit being converted. During calibration, the estimation D/A input code is set to the top M bits of $D_{RAM A}$, allowing the value of $X_{OD/A}$ to be adjusted with the same $X_{ED/A}$ as will be applied during conversion. However, when the reference D/A is being calibrated on the n th PWL segment, with $X_{in}:(n+1)2^L$ applied at the S/H, $D_{ED/A}$ is held at $n2^L$, since it would be set at that code when

TABLE II
MODEL EQUATIONS FOR THE A/D CONVERTER OF FIG. 5

$X_{SH}(X_{in}) = O_{SH} + A_{SH}X_{in} + f_{SH}(X_{in})$
$X_{MD/A}(D_{MD/A}, D_{RAM B}) = \frac{D_{MD/A}}{2^L} X_{RD/A} + O_{MD/A} + E_{MD/A}(D_{MD/A}, D_{RAM B})$ where $E_{MD/A}(0, D_{RAM B}) = E_{MD/A}(2^L, D_{RAM B}) = 0$ by definition
$X_{CD/A}(D_{CD/A}) = X_{in,min} + (X_{in,max} - X_{in,min}) \frac{D_{CD/A}}{2^N}$
$X_{OD/A}(D_{RAM A}=k)$ refers to the offset D/A output when its input is specified by the value stored in RAM A at location k .
$X_{RD/A}(D_{RAM B}=k)$ refers to the reference D/A output when its input is specified by the value stored in RAM B at location k .
$X_{amp}(X_{SH}, X_{ED/A}) = O_{amp} + A_{amp}(X_{SH} - X_{ED/A}) + f_{amp}(X_{SH}, X_{ED/A})$
$X_O = \begin{cases} 1, & X_{comp+} \geq O_{comp} \\ 0, & X_{comp+} < O_{comp} \end{cases}$
$X_{in:k}$ is shorthand for $X_{in} = X_{in,min} + (X_{in,max} - X_{in,min}) \frac{k}{2^N}$
$X_{ED/A:k}$ is shorthand for $X_{ED/A}(D_{ED/A} = k)$, which is the output of the estimation D/A with input = k .

converting within that segment. An analysis like that done on Fig. 4 was performed for this converter using the model and system equations shown in Table II. The resulting error budget equation is given by

$$\begin{aligned}
 & \frac{2^N}{A_{SH} A_{amp} (X_{in,max} - X_{in,min})} \\
 & \times \left\{ \max \left| A_{amp} [f_{SH}(X_{in} : n2^L) \right. \right. \\
 & + \frac{m}{2^L} [f_{SH}(X_{in} : [n+1]2^L) \\
 & - f_{SH}(X_{in} : n2^L)] - f_{SH}(X_{in} : n2^L + m)] \left| \right. \\
 & + \max \left| f_{amp}(X_{SH}(X_{in} : n2^L), X_{ED/A} : n2^L) \right. \\
 & + \frac{m}{2^L} [f_{amp}(X_{SH}(X_{in} : [n+1]2^L), X_{ED/A} : n2^L) \\
 & - f_{amp}(X_{SH}(X_{in} : n2^L), X_{ED/A} : n2^L)] \\
 & - f_{amp}(X_{SH}(X_{in} : n2^L + m), X_{ED/A} : n2^L) \left| \right. \\
 & + \max [|Q_{OD/A}(D_{RAM A})|, |Q_{RD/A}(D_{RAM B})|] \\
 & + \max |E_{MD/A}(D_{MD/A}, D_{RAM B})| \left. \right\} < \frac{1}{2}, \\
 & \forall n \in [0, 2^{N-L} - 1] \text{ for every } m \in [0, 2^L - 1] \quad (2)
 \end{aligned}$$

and is very similar to the one for Fig. 4. The main difference is now the nonlinearity from both the S/H and amplifier are included.

V. SIMULATIONS

A 16 b converter using the architecture of Fig. 5 was simulated behaviorally using the Saber simulator [22], and the various nonidealities considered in the analysis were included in the simulation. The PWL approximation was formed by

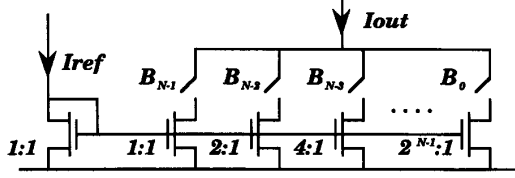


Fig. 6. Simple current-steering multiplying D/A converter. An actual implementation would likely use higher output resistance current sources, such as cascodes, and multiple mirror stages for a large number of bits.

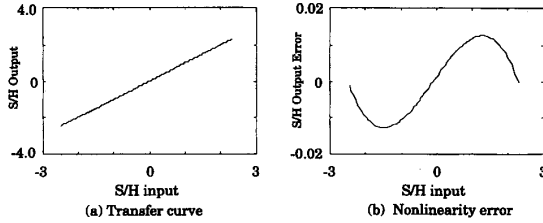


Fig. 7. Behavioral simulations of sample-and-hold characteristics. The output of the S/H was described mathematically as a function of the input, with offset, gain error, and nonlinearity added to model actual circuit nonlinearities.

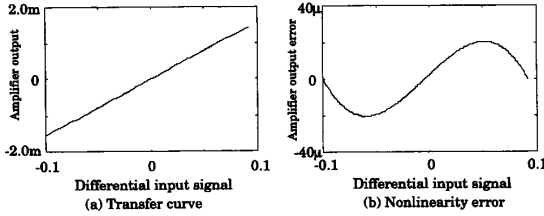


Fig. 8. Behavioral simulations of amplifier characteristics, based on a model generated like that of the S/H.

dividing the input region into 1 024 segments, requiring 1 024 offset and 1 024 slope calibrations. The S/H was modeled with 7 b linearity over the $[-2.5, +2.5]$ input range, and its transfer curve and nonlinearity error are shown in Fig. 7. Similarly, Fig. 8 shows the transfer function and nonlinearity error from X_{in} to the amplifier output with $X_{ED/A}$ set to zero. The combined amplifier and S/H characteristic had approximately 6 b linearity over the range $X_{in} \in [-0.1, +0.1]$, but the linearity error over each PWL segment was within $\pm 1/10$ LSB. Six bit converters were used for the estimation D/A and MD/A, each of which had 7 b accuracy. The offset and reference D/A's were constructed using 7 b accuracy internally, and extra bits were included for redundancy to ensure sufficient adjustability. The 18 b offset D/A covered a range of ± 2048 LSB's, with $1/16$ LSB adjustability and two bits used for redundancy. The 10 b reference D/A used one bit for redundancy and had a range of ± 32 LSB's with $1/8$ LSB adjustability.

Using Saber, the converter was simulated as going through a full calibration phase. Approximately 30 h of simulation time was required to calibrate the full converter on a Sun 3/260, and 28.7 kb of simulated RAM was needed to hold the calibration data. Evaluating the accuracy of the calibrated converter requires determining the switching point input level for every digital output code and can be done by sweeping the converter

input in fractions of an LSB and noting when the digital output code changes state. Determining each switching point to within $1/8$ LSB accuracy was estimated to require approximately one year of simulation time on the same platform and so was not performed. Instead, the input was stepped in $1/8$ LSB increments over short ranges of 25 LSB's at four points along the total A/D input range. The differential nonlinearity (DNL) and integral nonlinearity (INL) of the converter over these regions is shown in Fig. 9. The integral nonlinearity plots were obtained using the same ideal straight line approximation for all cases, not one optimized for each region.

In most successive-approximation A/D converters, the worst case error generally occurs at major bit transitions, since that is usually where the internal D/A converter has its worst errors. However, the simulated converter has calibration points at all these major bit transitions, preventing the errors there from being outstanding, as is seen in Fig. 9(c) and (d) at code 32 768 ($1/2$ full scale). Instead, the converter's worst case errors tend to occur where the MD/A has its worst case error, which can be seen in Fig. 10 to be at $1/4$, $1/2$, and $3/4$ full scale. This corresponds to the error in Fig. 9(e) and (f) at code 32 799, which is when the MD/A is right at $1/2$ full scale. Similarly, the worst error in Fig. 9(g) and (h) occurs when the MD/A is right at $3/4$ full scale. These results indicate that the converter nonlinearity lies approximately within ± 0.4 LSB's if the rest of the converter range can be assumed to be no worse than these small segments. This corresponds well with the result using (2), which predicted a worst possible error of 0.48 LSB's.

Although the converter architectures discussed thus far can tolerate various circuit nonlinearities, there are still other potential circuit problems that are not addressed. Noise is a primary consideration in a high resolution converter, including both electronic noise and the coupling of digital switching transients into sensitive analog circuitry. Although the former has not changed significantly from the case of a typical successive approximation converter, the latter may be a serious concern due to the presence of the large RAM on chip. The desired level of resolution and speed of conversion will determine the degree of isolation necessary between the analog and digital circuitry to reduce coupling through signal lines, power supplies, and the IC substrate.

VI. IMPLEMENTATION CONSIDERATIONS

The circuit implementation of a converter like that of Fig. 5 can be designed in a variety of ways, depending on the particular application. The most critical component in the system is the calibration D/A, which will limit the resolution attainable from the entire converter no matter how much memory is used. However, since this component need not operate at high speed, a relatively slow design may be used, such as a delta-sigma based architecture or an integrating ramp topology. If a voltage input is required, which is a common scenario, then the S/H and estimation D/A would most likely be voltage-mode as well. The speed of the converter can be increased significantly by putting a low resolution flash A/D at the S/H output and using its output for the estimation D/A

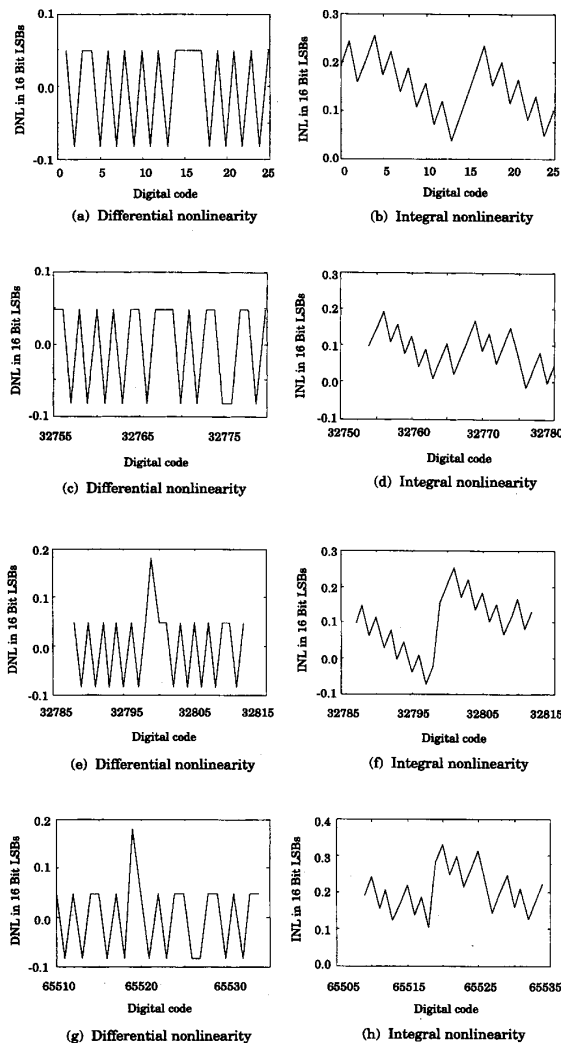


Fig. 9. Differential and integral nonlinearity of the simulated 16 b A/D converter are shown. Due to long simulation time, the A/D was only simulated enough to extract DNL and INL information over several small regions of the total input range, which are shown above for the corresponding digital output code regions.

input. If the flash converter provides enough resolution, then the estimation D/A will not need to be changed again during the conversion. In this case, the amplifier must operate only once per conversion, further increasing speed.

Since the amplifier need not be highly linear over its entire input range, an open-loop design can be used to provide fast settling. The summation before the comparator can be implemented using switched-capacitor techniques, although a current-mode sum would be much more area-efficient. Taking the current sum approach requires a transconductance amplifier for the amplifier and a current-mode offset D/A and MD/A. Fig. 6 shows a simplified current-steering MD/A with a current reference, which then requires a current-mode reference D/A. The offset D/A and reference D/A are easily realized using a similar architecture with a constant reference current. The

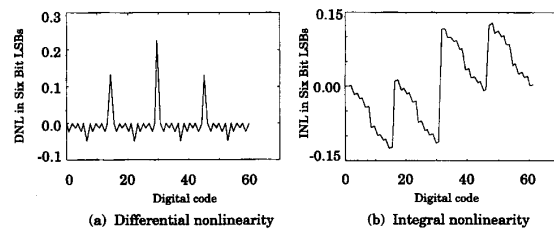


Fig. 10. Nonlinearity of 6 b MD/A.

design of the remaining analog block, the comparator, is fairly straightforward, requiring a current input and yielding a digital voltage level output for interfacing with the digital logic.

VII. CONCLUSION

In the arena of A/D converter design, there exists a tradeoff between attainable speed and resolution which is readily apparent. High speed converters can be built at the cost of low resolution, while high resolution can be achieved if speed is sacrificed. The techniques discussed here are intended to draw from the best of both approaches to yield both high speed and high resolution in a converter. The example successive approximation architectures shown here were chosen for their simplicity in order to best demonstrate the nonlinearity correction techniques presented. However, these techniques can be best exploited for high conversion speed in a more parallel converter architecture, such as a multistage flash or a pipeline configuration. Even in cases where lower precision specifications do not lead to faster designs, they simplify the design process and may still result in power or die area savings.

There is a primary difference in the design requirements of circuitry for a nonlinearity correcting A/D and a conventional architecture. In most conventional converters, the circuit designs must i) meet the necessary resolution in the static case, and ii) maximize the speed at which dynamic performance deviates significantly from this static performance. In the nonlinearity correcting architecture, these requirements are separated between the calibration circuitry and the main path circuitry. The calibration circuit design must meet the necessary resolution, but its speed performance is not critical. The main path circuit design must maximize the speed at which dynamic performance deviates from static performance, but the static performance need not have high resolution. By restricting highly linear, highly accurate circuitry to a calibration function, the conversion speed is then determined primarily by less accurate but faster circuitry that is periodically recalibrated to maintain resolution. Similarly, the use of a high speed but potentially nonlinear S/H can lead to a high analog input bandwidth limited by the jitter and slew of digital control signals. This is particularly important for current-mode A/D's, where highly linear S/H's are difficult to build at high speed.

APPENDIX

In order to calibrate the A/D converter of Fig. 4, the offset and slope of each PWL segment must be adjusted to match

the S/H output curve. Using L bits in the MD/A means 2^{N-L} segments are used, each having length 2^L LSB's. During the offset calibration, the digital codes $D_{CD/A} = n2^L$ are applied to the calibration D/A for each $n \in [0, 2^{N-L} - 1]$. For each of the codes, $D_{MD/A}$ and $D_{RD/A}$ are set to zero, and a binary search is performed on $D_{OD/A}$ using X_0 until $X_{comp+} \approx O_{comp}$. The final value of $D_{OD/A}$ is stored in RAM A at the address $D_{RAM A} = n$. The final value of $X_{OD/A}$ is then given by

$$\begin{aligned} X_{OD/A}(D_{RAM A} = n) \\ = X_{SH}(X_{in} : n2^L) \\ - X_{MD/A}(0, D_{RAM B}) - O_{comp} \\ + Q_{OD/A}(D_{RAM A} = n), n \in [0, 2^{N-L} - 1] \end{aligned} \quad (3)$$

where $Q_{OD/A}(D_{RAM A} = n)$ is the error at the output of the offset D/A due to its finite resolution (i.e., quantization error). $X_{in} : k$ is shorthand notation for

$$X_{in} : k = X_{in,min} + (X_{in,max} - X_{in,min})k2^{-N}, k \in [0, 2^N]. \quad (4)$$

The value of $D_{RAM B}$ in (3) is not important, since $X_{MD/A}$ is independent of $D_{RAM B}$ when $D_{MD/A} = 0$.

When calibrating the slopes of the PWL segments, the digital codes $D_{CD/A} = (n+1)2^L$ are applied to the calibration D/A for each $n \in [0, 2^{N-L} - 1]$. Note that these codes are identical to those needed for the offset calibration, except for the first and last, and so both calibrations can be combined. The final code $D_{CD/A} = 2^N$ is achieved by adding an extra LSB to the calibration D/A output, as in the case of the MD/A. For each code $D_{CD/A} = (n+1)2^L$, $D_{RAM A}$ is set to n and $D_{MD/A}$ is set to 2^L . A binary search is then performed using $D_{RD/A}$ until the rightmost endpoint of the PWL segment matches the S/H output curve. The final value of $D_{RD/A}$ is then stored in RAM B at address $D_{RAM B} = n$, and the final value of X_{ref} is given by

$$\begin{aligned} X_{ref}(D_{RAM B} = n) \\ = X_{SH}(X_{in} : [n+1]2^L) - X_{OD/A}(D_{RAM A} = n) \\ - X_{MD/A}(D_{MD/A} = 0, D_{RAM B} = n) - O_{comp} \\ + Q_{RD/A}(D_{RAM B} = n), \forall n \in [0, 2^{N-L} - 1] \end{aligned} \quad (5)$$

where $Q_{RD/A}(D_{RAM B})$ is the quantization error of the reference D/A.

The error budget for the converter can be calculated by solving for the value of X_{in} at which the converter output switches from one digital code to the next higher code. By forcing each switching point to fall within ± 0.5 LSB of its ideal value, the A/D transfer curve will be bound to within a 1 LSB wide band over its entire range. Using $X_{in} : i$ as the ideal switching point between codes $D_{out} = i-1$ and $D_{out} = i$, the error budget equation will be of the form

$$X_{in} : (i-1/2) < X_{in,sw,i} < X_{in} : (i+1/2), \forall i \in [1, 2^N - 1] \quad (6)$$

where $X_{in,sw,i}$ is the actual value of X_{in} at which D_{out} switches from $i-1$ to i . However, it is possible to translate the

equation through the S/H function and simplify the resulting analysis, yielding

$$\begin{aligned} X_{SH}(X_{in} : [i-1/2]) < X_{SH}(X_{in,sw,i}) \\ < X_{SH}(X_{in} : [i+1/2]), \forall i \in [1, 2^N - 1]. \end{aligned} \quad (7)$$

The switching from output code $i-1$ to i occurs when

$$\begin{aligned} X_{SH}(X_{in,sw,i}) - X_{OD/A}(D_{RAM A}) \\ - X_{MD/A}(D_{MD/A}, D_{RAM B}) = O_{comp}, \end{aligned} \quad (8)$$

or, rewriting (8),

$$\begin{aligned} X_{SH}(X_{in,sw,i}) = X_{OD/A}(D_{RAM A}) \\ + X_{MD/A}(D_{MD/A}, D_{RAM B}) + O_{comp}. \end{aligned} \quad (9)$$

Looking first at the major bit transitions, the switching points are given by

$$\begin{aligned} X_{SH}(X_{in,sw,n2^L}) = X_{OD/A}(D_{RAM A} = n) \\ + X_{MD/A}(D_{MD/A} = 0, D_{RAM B}) \\ + O_{comp}, \forall n \in [1, 2^{N-L} - 1] \end{aligned} \quad (10)$$

which can be simplified using (3), yielding

$$\begin{aligned} X_{SH}(X_{in,sw,n2^L}) = X_{SH}(X_{in} : n2^L) \\ + Q_{OD/A}(D_{RAM A} = n), \\ \forall n \in [1, 2^{N-L} - 1]. \end{aligned} \quad (11)$$

The equation for the switching points between major bit transitions is very similar and is given by

$$\begin{aligned} X_{SH}(X_{in,sw,n2^L+m}) \\ = X_{OD/A}(D_{RAM A} = n) \\ + X_{MD/A}(D_{MD/A} = m, D_{RAM B} = n) \\ + O_{comp}, \forall n \in [0, 2^{N-L} - 1] \\ \text{for every } m \in [1, 2^L - 1]. \end{aligned} \quad (12)$$

Using (3) and Table I, this equation can be reduced further to

$$\begin{aligned} X_{SH}(X_{in,sw,n2^L+m}) \\ = X_{SH}(X_{in} : n2^L) + Q_{OD/A}(D_{RAM A} = n) \\ + E_{MD/A}(D_{MD/A} = m, D_{RAM B} = n) \\ + m2^{-L}X_{ref}, \forall n \in [0, 2^{N-L} - 1] \\ \text{for every } m \in [1, 2^L - 1] \end{aligned} \quad (13)$$

where $E_{MD/A}$ is the error of the MD/A output to a straight line. X_{ref} in (13) can be expanded using (3) and (5), and further simplification yields

$$\begin{aligned} X_{SH}(X_{in,sw,n2^L+m}) \\ = X_{SH}(X_{in} : n2^L) + \frac{m}{2^L}[X_{SH}(X_{in} : [n+1]2^L) \\ - X_{SH}(X_{in} : n2^L)] + \left(1 - \frac{m}{2^L}\right)Q_{OD/A}(D_{RAM A} = n) \\ + \frac{m}{2^L}Q_{RD/A}(D_{RAM B} = n) \\ + E_{MD/A}(D_{MD/A} = m, D_{RAM B} = n), \\ \forall n \in [0, 2^{N-L} - 1] \\ \text{for every } m \in [1, 2^L - 1]. \end{aligned} \quad (14)$$

Note that by setting $m = 0$, (14) reduces to (11), so both cases are contained in (14). Now this expression for $X_{SH}(X_{in,sw,n2^L+m})$ can be substituted into the error budget

(7), and expanding all $X_{SH}(X_{in})$ terms into their offset, gain, and nonlinearity components gives, after simplifying,

$$\begin{aligned}
 & A_{SH}(X_{in,max} - X_{in,min}) \left(-\frac{1}{2}\right) 2^{-N} \\
 & + f_{SH}\left(X_{in} : \left[n2^L + m - \frac{1}{2}\right]\right) < f_{SH}(X_{in} : n2^L) \\
 & + \frac{m}{2^L} [f_{SH}(X_{in} : [n+1]2^L) - f_{SH}(X_{in} : n2^L)] \\
 & + \left(1 - \frac{m}{2^L}\right) Q_{OD/A}(D_{RAM A} = n) \\
 & + \frac{m}{2^L} Q_{RD/A}(D_{RAM B} = n) \\
 & + E_{MD/A}(D_{MD/A} = m, D_{RAM B} = n) \\
 & < A_{SH}(X_{in,max} - X_{in,min}) \left(\frac{1}{2}\right) 2^{-N} \\
 & + f_{SH}\left(X_{in} : \left[n2^L + m + \frac{1}{2}\right]\right), \\
 & \forall n \in [0, 2^{N-L} - 1] \text{ for every } m \in [0, 2^L - 1]. \quad (15)
 \end{aligned}$$

If the S/H nonlinearity is smooth, it is reasonable to assume that the nonlinearity does not change significantly over 1 LSB, so that

$$\begin{aligned}
 & f_{SH}\left(X_{in} : n2^L + m - \frac{1}{2}\right) \\
 & \approx f_{SH}(X_{in} : n2^L + m) \approx f_{SH}\left(X_{in} : n2^L + m + \frac{1}{2}\right), \\
 & \forall n \in [0, 2^{N-L} - 1] \text{ for every } m \in [0, 2^L - 1]. \quad (16)
 \end{aligned}$$

Using this assumption in (15) leads to

$$\begin{aligned}
 & -\frac{1}{2} < \frac{2^N}{A_{SH}(X_{in,max} - X_{in,min})} \left\{ f_{SH}(X_{in} : n2^L) \right. \\
 & + \frac{m}{2^L} [f_{SH}(X_{in} : [n+1]2^L) - f_{SH}(X_{in} : n2^L)] \\
 & - f_{SH}(X_{in} : n2^L + m) \\
 & + \left(1 - \frac{m}{2^L}\right) Q_{OD/A}(D_{RAM A} = n) \\
 & + \frac{m}{2^L} Q_{RD/A}(D_{RAM B} = n) \\
 & + E_{MD/A}(D_{MD/A} = m, D_{RAM B} = n) \left. \right\} < +\frac{1}{2}, \\
 & \forall n \in [0, 2^{N-L} - 1] \text{ for every } m \in [0, 2^L - 1] \quad (17)
 \end{aligned}$$

in which the values have been normalized to units of LSB's by dividing through by $A_{SH}(X_{in,max} - X_{in,min})2^{-N}$, the average LSB size at the S/H output. This equation can be made more restrictive as

$$\begin{aligned}
 & \frac{2^N}{A_{SH}(X_{in,max} - X_{in,min})} \left\{ \max \left| f_{SH}(X_{in} : n2^L) \right. \right. \\
 & + \frac{m}{2^L} [f_{SH}(X_{in} : [n+1]2^L) - f_{SH}(X_{in} : n2^L)] \\
 & - f_{SH}(X_{in} : n2^L + m) \left. \right| \\
 & + \max[|Q_{OD/A}(D_{RAM A})|, |Q_{RD/A}(D_{RAM B})|] \\
 & + \max|E_{MD/A}(D_{MD/A}, D_{RAM B})| \left. \right\} < +\frac{1}{2}, \\
 & \forall n \in [0, 2^{N-L} - 1] \text{ for every } m \in [0, 2^L - 1]. \quad (18)
 \end{aligned}$$

Note that the quantization errors of the offset and reference

D/A's are combined together in the second maximum term. This is due to their weighting in (17) by $(1 - m2^{-L})$ and $m2^{-L}$, respectively, so that they each have full effect at opposite ends of each segment and are never summed when both are at their maximum.

The derivation of the error budget for the converter in Fig. 5 was done in similar fashion using the equations from Table II, yielding the final error budget given in (2).

ACKNOWLEDGMENT

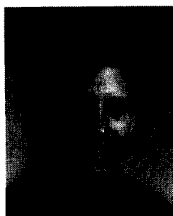
The authors would like to thank the anonymous reviewers for their comments and suggestions on improving the manuscript; their input was greatly appreciated.

REFERENCES

- [1] H. S. Lee and D. A. Hodges, "Self-calibration technique for A/D converters," *IEEE Trans. Circuits Syst.*, pp. 188-190, Mar. 1983.
- [2] H. S. Lee, D. A. Hodges, and P. R. Gray, "A self-calibrating 15 bit CMOS A/D converter," *IEEE J. Solid-State Circuits*, pp. 813-819, Dec. 1984.
- [3] J. C. Candy, "A use of double integration in sigma delta modulation," *IEEE Trans. Commun.*, pp. 249-258, Mar. 1985.
- [4] M. W. Hauser, P. J. Hurst, and R. W. Broderson, "MOS ADC-filter combination that does not require precision analog components," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, pp. 80-81, Feb. 1985.
- [5] B. Del Signore, D. Kerth, N. Sooch, and E. Swanson, "A monolithic 20 b delta-sigma A/D converter," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, pp. 170-171, Feb. 1990.
- [6] J. Fernandes, S. R. Lewis, A. M. Mallinson, and G. A. Miller, "A 14-bit 10-ms subranging A/D converter with S/H," *IEEE J. Solid-State Circuits*, pp. 1309-1315, Dec. 1988.
- [7] M. Nayeibi and B. A. Wooley, "A 10-bit video BiCMOS track-and-hold amplifier," *IEEE J. Solid-State Circuits*, pp. 1507-1516, Dec. 1989.
- [8] P. J. Lim and B. A. Wooley, "A high-speed sample-and-hold technique using a Miller hold capacitance," *IEEE J. Solid-State Circuits*, pp. 643-651, Apr. 1991.
- [9] F. Moraveji, "A high-speed current-multiplexed sample-and-hold amplifier with low hold step," *IEEE J. Solid-State Circuits*, pp. 1800-1808, Dec. 1991.
- [10] H. Ohara, H. X. Ngo, M. J. Armstrong, C. F. Rahim, and P. R. Gray, "A CMOS programmable self-calibrating 13-bit eight-channel data acquisition peripheral," *IEEE J. Solid-State Circuits*, pp. 930-938, Dec. 1987.
- [11] H. S. Lee and D. A. Hodges, "Accuracy considerations in self-calibrating A/D converters," *IEEE Trans. Circuits Syst.*, pp. 590-597, June 1985.
- [12] H. J. Schouwenaars, E. C. Dijkman, B. M. J. Kup, and E. J. M. Van Tuijl, "A monolithic dual 16-bit D/A converter," *IEEE J. Solid-State Circuits*, pp. 424-429, June 1986.
- [13] P. J. A. Naus, E. C. Dijkman, E. F. Stikvoort, D. J. Holland, and W. Bradinal, "A CMOS stereo 16-bit D/A converter for digital audio," *IEEE J. Solid-State Circuits*, pp. 390-395, June 1987.
- [14] H. J. Schouwenaars, D. W. J. Groeneveld, and H. A. H. Termeer, "A low-power stereo 16-bit CMOS D/A converter for digital audio," *IEEE J. Solid-State Circuits*, pp. 1290-7, Dec. 1988.
- [15] L. R. Carley, "A noise-shaping coder topology for 15+ bit converters," *IEEE J. Solid-State Circuits*, pp. 267-273, Apr. 1989.
- [16] Y. Manoli, "A self-calibration method for fast high-resolution A/D and D/A converters," *IEEE J. Solid-State Circuits*, pp. 603-608, June 1989.
- [17] Y. Matsuya, K. Uchimura, A. Iwata, and T. Kaneko, "A 17 bit oversampling D-A conversion technology using multistage noise shaping," *IEEE J. Solid-State Circuits*, pp. 969-975, Aug. 1989.
- [18] D. W. J. Groeneveld, H. J. Schouwenaars, H. A. H. Termeer, and C. A. A. Bastiaansen, "A self-calibration technique for monolithic high-resolution D/A converters," *IEEE J. Solid-State Circuits*, pp. 1517-1522, Dec. 1989.
- [19] B. M. J. Kup, E. C. Dijkman, P. J. A. Naus, and J. Snee, "A bit-stream digital-to-analog converter with 18-b resolution," *IEEE J. Solid-State Circuits*, pp. 1757-1763, Dec. 1991.
- [20] H. J. Schouwenaars, D. W. J. Groeneveld, C. A. A. Bastiaansen, and H. A. H. Termeer, "An oversampled multibit CMOS D/A converter for

digital audio with 115-dB dynamic range," *IEEE J. Solid-State Circuits*, pp. 1775-1780, Dec. 1991.

- [21] K.-S. Tan, S. Kiriaki, M. de Wit, J. W. Fattaruso, C.-Y. Tsay, W. E. Matthews, and R. K. Hester, "Error correction techniques for high-performance differential A/D Converters," *IEEE J. Solid-State Circuits*, vol. 25, no. 6, pp. 1318-1327, Dec. 1990.
- [22] Electronics staff, "SABER cuts SPICE out of analog simulation," *Electronics*, pp. 80-82, Oct. 30, 1986.



Terry L. Sculley (S'87-M'92) attended the University of Florida and the Georgia Institute of Technology, where he received the B.E.E. (Highest Honors), M.S.E.E., and Ph.D. degrees in electrical engineering in 1987, 1990, and 1992, respectively.

He was the recipient of a Semiconductor Research Corporation Graduate Fellowship and a Georgia Tech President's Fellowship during his graduate studies. Since 1992 he has been with Washington State University, where he is an Assistant Professor with the School of Electrical Engineering and

Computer Science. His current research interests are in analog and mixed analog/digital integrated circuits, systems, and signal processing, and he is presently an investigator with the National Science Foundation Center for Design of Analog/Digital Integrated Circuits.

Dr. Sculley is a member of Tau Beta Pi and Eta Kappa Nu.



Martin A. Brooke (S'85-M'87) received the B.E. (Elect.) degree (first class Hons.) from Auckland University, New Zealand in 1981. He received the M.S. and Ph.D. degrees in electrical engineering from the University of Southern California in 1984 and 1988, respectively. His doctoral research focused on reconfigurable analog and digital integrated circuit design.

He is currently an Analog Devices Inc. Career Development Associate Professor of Electrical Engineering at the Georgia Institute of Technology.

Professor Brooke won a 1990 National Science Foundation Research Initiation Award in the analog signal processing area, was awarded the 1990 Georgia Tech Sigma Xi Research Award for Undergraduate Research Advisor, and 1990 Georgia Tech Undergraduate Research Opportunity Program, Outstanding Research Award, and the 1992 IEEE Midwest Symposium on Circuits and Systems, Myril B. Reed Best Paper Award. He has one patent awarded, and four are pending. He has published 24 refereed papers in journals and conference proceedings, and articles on his work have appeared in 15 different trade publications. He has developed two software programs for analog CAD. He is a member of the INNS.