

A Background Comparator Calibration Technique for Flash Analog-to-Digital Converters

Chun-Cheng Huang, *Student Member, IEEE* and Jieh-Tsong Wu, *Member, IEEE*

Abstract—This paper presents a background calibration technique for trimming the input-referred offsets of the comparators in a flash analog-to-digital converter (ADC) without interrupting the ADC's normal operation. For a random-chopping comparator, the polarity of its offset is detected by observing the code density of its comparison results. Binary feedback is then used to digitally adjust the comparator's offset so that the offset is minimized. All calibration procedures are performed in the digital domain. The calibration performance is characterized by the converging speed of the feedback loop and the offset fluctuation due to the disturbance of the ADC's input. These two performance indexes of a background-calibrated comparator (BCC) are determined by three parameters: the probabilistic distribution of the ADC's input, the BCC's offset quantized step size, and the threshold of an internal bilateral peak detector. The offset fluctuation of a BCC can be drastically reduced by employing a windowing mechanism. The use of windowed BCCs in a flash ADC can introduce nonmonotonic-threshold (NMT) effects which include an increase in calibration settling time and an increase in $\sigma(V_{OS})$. The use of uncorrelated random chopping for neighboring BCCs can ensure the validity of offset detection and mitigate the NMT effects.

Index Terms—Comparator, flash analog-to-digital converter (ADC), offset calibration.

I. INTRODUCTION

SHOWN in Fig. 1 is an N -bit flash analog-to-digital converter (ADC) that uses $2^N - 1$ comparators to simultaneously compare input, V_i , with $2^N - 1$ references, $V_{R,j}$, where $j = 1, 2, \dots, 2^N - 1$. The overall digital output D_o is obtained by encoding the binary outputs from the comparators. The flash architecture has the highest analog-to-digital (A/D) conversion speed at a given N for a given technology, since it does not require linear amplification. For ADCs with large N , circuit techniques, such as subranging, folding, and interpolation, have been used to reduce the size of the comparator array, so that the power consumption and the total input capacitive loading are reduced [1], [2].

For a high-speed CMOS flash ADC, the linearity of its transfer function is predominantly degraded by the random input-referred offset voltages of the comparators. The offset of a comparator with symmetric circuit configuration is caused by device mismatches. Devices with larger size have better

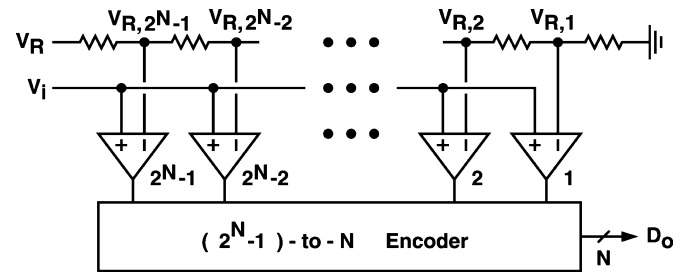


Fig. 1. Conventional flash ADC architecture.

matching properties but also result in circuits with less power efficiency. Due to this design consideration for matching, there exists a fundamental tradeoff among the speed, power, and accuracy for a CMOS flash ADC [3].

To overcome this inherent device's constraint, several techniques have been proposed, such as switched-capacitor offset cancellation [4], circuit-level spatial filtering [5], digitally controlled offset trimming [6], [7], and calibrated redundancy [8]. The last two schemes are potentially more power efficient than other techniques, since no extra clock phases are required and no extra circuitry to consume power. However, both of them require calibration mechanism to detect the offsets and adjust the circuit configuration accordingly. In these cases, foreground offset calibration schemes, which can only be executed once for a specific time period, may not be sufficient to prevent offsets shift due to the variation of temperature and supply voltage.

This paper describes a comparator calibration technique that can perform offset trimming in the background without interrupting its normal comparison operation. Since most of the required circuit overhead for the proposed scheme is in the digital domain and little modification is done to the analog critical signal path, the proposed scheme will not degrade the speed of the circuit's comparison function. The technique can be applied simultaneously to all comparators in a flash ADC to improve its linearity.

The rest of this paper is organized as follows. Section II describes a random-chopping comparator (RCC) and its probabilistic characteristics. Section III gives the design and analysis of a background-calibrated comparator (BCC) based on the random-chopping technique. Section IV gives the design and analysis of a flash ADC using the BCCs. Related design issues are also discussed. Section V draws conclusions. Finally, the Appendix includes a mathematical treatment for the offset fluctuation behavior of the proposed BCC.

In the following analysis, a 6-bit flash ADC is used as a design example. For this ADC, its input range is $\pm V_{FS}/2$, and one LSB

Manuscript received September 17, 2004; revised January 11, 2005. This work was supported by the National Science Council of Taiwan, R.O.C. under Contract NSC-93-2220-E-009-005, and the MediaTek Research Center at National Chiao-Tung University. This paper was recommended by Associate Editor A. Wang.

The authors are with the Department of Electronics Engineering, National Chiao-Tung University, Hsin-Chu, Taiwan, R.O.C. (e-mail: jtwu@mail.nctu.edu.tw).

Digital Object Identifier 10.1109/TCSI.2005.852198

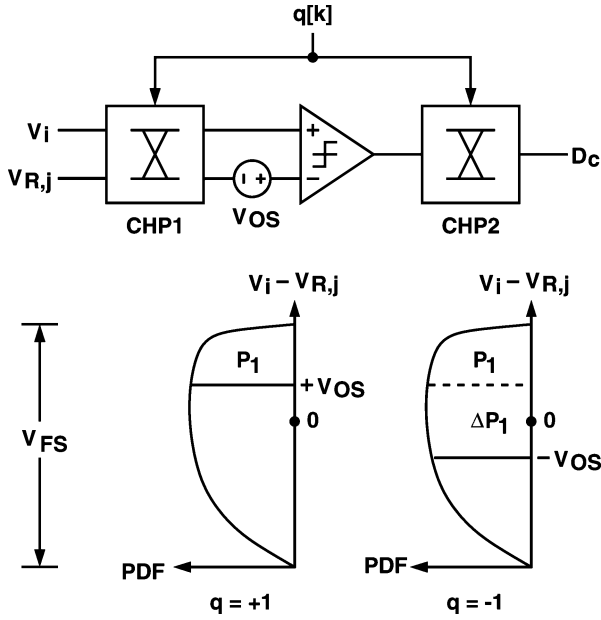


Fig. 2. Comparator with random choppers.

is defined as $V_{FS}/2^6 = V_{FS}/64$. The ADC's input is assumed to be a full-scale sinusoidal signal, i.e., $V_i = (V_{FS}/2) \sin(\omega_i t)$.

II. RANDOM-CHOPPING COMPARATOR

The proposed comparator calibration scheme is based on the RCC shown in Fig. 2. This chopping comparator can replace the j th comparator shown in Fig. 1. The comparator compares input V_i , with the j th reference voltage $V_{R,j}$, and then generates a corresponding binary output $D_c[k] \in \{1, 0\}$. Due to the clocked operation of the comparator, $D_c[k]$ is a discrete signal with k indicating the discrete time index. The internal comparator has an input-referred offset voltage of V_{OS} . The two choppers CHP1 and CHP2 are controlled by a binary-valued random sequence, $q[k] \in \{+1, -1\}$. CHP1 is an analog chopper, which passes the inputs unchanged when $q = +1$ and interchanges the inputs when $q = -1$. In CMOS technologies, CHP1 can be realized using 4 analog switches. CHP2 is a digital chopper, which inverts the digital comparison result when $q = -1$.

Also shown in Fig. 2 is the probability density function (PDF) of $V_i - V_{R,j}$. When $q = +1$, the probability for $D_c = +1$ is P_1 . When $q = -1$, the probability for $D_c = +1$ is $P_1 + \Delta P_1$. Thus, one can detect the polarity of V_{OS} from the polarity of ΔP_1 , and then trim the V_{OS} accordingly. It is necessary for the random sequence $q[k]$ to be uncorrelated with V_i , so that the comparator can perceive identical PDF of $V_i - V_{R,j}$, regardless of $q[k]$ being $+1$ or -1 .

The comparator's V_{OS} can be adjusted by reconfiguring its low-speed section, which is separated from the high-speed signal path [6], [7], [9]. Thus, this added V_{OS} controllability costs little speed/power penalty.

Random choppers have been used to extract ADC's input offset [10], [11]. But when applied to comparators, only the polarity of the V_{OS} can be detected, since the exact PDF of V_i is not available. Due to its highly nonlinear characteristic, we use probabilistic system analysis techniques to analyze its behavior.

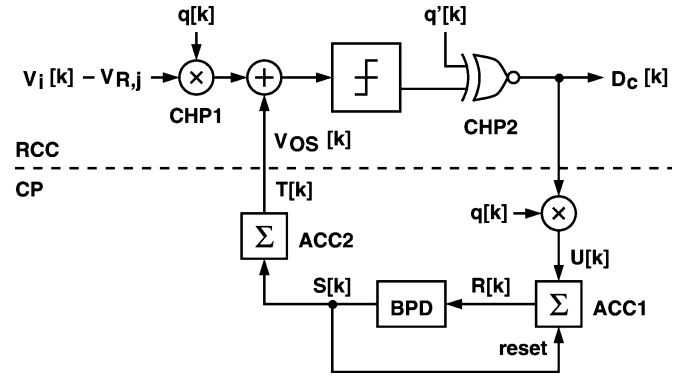


Fig. 3. BCC including RCC and CP.

III. BACKGROUND-CALIBRATED COMPARATOR

Fig. 3 shows the block diagram of the proposed BCC, which is composed of an RCC and a calibration processor (CP). The CP resembles a discrete-time integrator in the digital domain [12], [13]. The CHP2 chopper in Fig. 2 becomes an XNOR gate controlled by a random sequence $q'[k] \in \{1, 0\}$ with signal pattern identical to that of $q[k]$. If $q[k] = +1$, $q'[k] = +1$; if $q[k] = -1$, $q'[k] = 0$. The ACC1 accumulator records the difference between the number of $D_c[k] = 1$ occurrences for $q[k] = +1$ and $q[k] = -1$. The ACC1's output is $R[k]$. The rate of long-term change in $R[k]$ is proportional to the probability difference ΔP_1 of Fig. 2. The bilateral peak detector (BPD) monitors the value of $R[k]$ and generates a corresponding triple-valued output, $S[k] \in \{+1, 0, -1\}$. The BPD has two thresholds, $+N_C$ and $-N_C$. When $R[k] > +N_C$, $S[k] = +1$. When $R[k] < -N_C$, $S[k] = -1$. Otherwise, $S[k] = 0$. In addition, if $S[k] = +1$ or $S[k] = -1$, the ACC1 accumulator will be reset in the following clock cycle. Thus, $-(N_C + 1) \leq R[k] \leq +(N_C + 1)$, and $S[k]$ can only remain as $+1$ or -1 for one clock cycle. The $S[k]$ sequence is integrated by the ACC2 accumulator. Its output, $T[k]$, controls the comparator's input offset voltage. The time-varying offset voltage can be expressed as

$$V_{OS}[k] = V_0 + \Delta V \times T[k] \quad (1)$$

where V_0 is the inherent comparator offset when $T[k] = 0$, and ΔV is the step size of the offset control.

Circuit realization of the CP is straightforward. No multi-bit multiplier is required. The BPD's threshold, N_C , can be chosen to be the power of 2 so that digital comparators are not required in the BPD. Then, the entire CP consists of only two accumulators and one XNOR gate.

The proposed calibration scheme does not require the exact information of V_i 's PDF. However, the effectiveness of the calibration depends on the probability difference ΔP_1 of Fig. 2. When the V_{OS} is large, the ΔP_1 is also large, the BPD get activated often, and the V_{OS} moves fast toward zero. When the V_{OS} is close to zero, the ΔP_1 also becomes small, then the BPD rarely get activated, and the V_{OS} is more stationary. Compared to the periodical adjusting scheme of [10], the proposed scheme has faster converging speed and lower offset fluctuation.

There are two design parameters in this calibration scheme, ΔV and N_C . Both parameters affect the converging speed as well as the offset fluctuation due to the disturbance of the input.

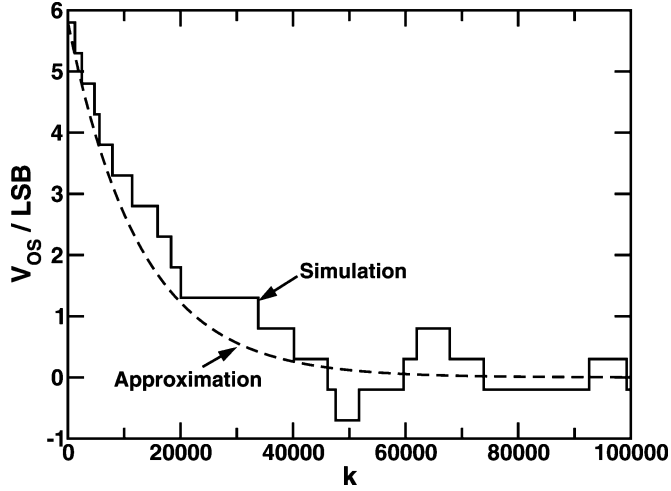


Fig. 4. $V_{OS}[k]$ transient response of a BCC example. $\Delta V = (1/2)$ LSB and $N_C = 64$. The initial condition is $V_{OS}[0] = 5.8$ LSB.

Large ΔV and small N_C result in fast converging speed but large fluctuation in the V_{OS} . On the other hand, small ΔV and large N_C result in small V_{OS} fluctuation but also slow converging speed. Detailed analyses are given in the following subsections.

A. Transient Behavior

In a BCC, its CP forces the RCC's internal V_{OS} moving toward zero with the mechanism of feedback. If the V_i 's PDF over the range of $V_{R,j} \pm V_{OS}$ is a constant $D(V_{R,j})$, the value of ΔP_1 of Fig. 2 can be expressed as $2D(V_{R,j})V_{OS}$. Then, the transient behavior of the feedback system can be approximated by a single-pole model

$$\frac{dV_{OS}[k]}{dk} = -\Delta V \times D(V_{R,j})V_{OS}[k] \times \frac{1}{N_C}. \quad (2)$$

Equation (2) is obtained by observing that the V_{OS} is changed by one ΔV only after the V_i occurs in the ΔP_1 region for $2N_C$ samples, i.e., N_C samples during $q[k] = +1$ and N_C samples during $q[k] = -1$. Thus, the averaged number of input samples required to change V_{OS} by one ΔV is $2N_C/|\Delta P_1| = N_C/[D(V_{R,j})|V_{OS}|]$. From (2), transient response of $V_{OS}[k]$ can be expressed as

$$V_{OS}[k] = V_{OS}[0] \cdot \exp\left[-\frac{k}{\tau_c}\right] \quad (3)$$

where the time constant $\tau_c = N_C/[\Delta V \cdot D(V_{R,j})]$. Shorter τ_c also results in better tracking ability for the calibration loop against environmental changes, such as temperature and supply voltage variation.

Fig. 4 shows an example of the $V_{OS}[k]$ transient response of a BCC. The initial offset, $V_{OS}[0]$, is set at 5.8 LSB. Calibration design parameters are $\Delta V = (1/2)$ LSB and $N_C = 64$. The BCC is assumed to be located at the middle of the 6-bit ADC design case mentioned in Section I. With a full-range sinusoidal ADC input and $V_{R,j} = 0$, we have $D(V_{R,j}) = (2/\pi) \cdot (1/V_{FS})$. In Fig. 4, the solid line is the discrete-time simulation result,

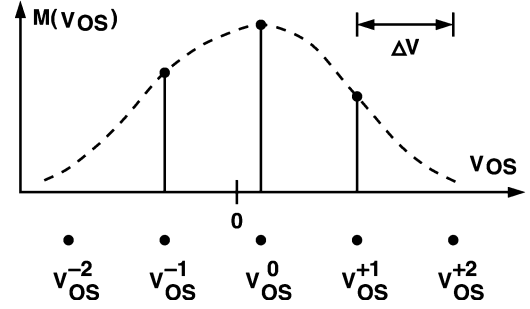


Fig. 5. Probability mass function of V_{OS} , $M(V_{OS})$.

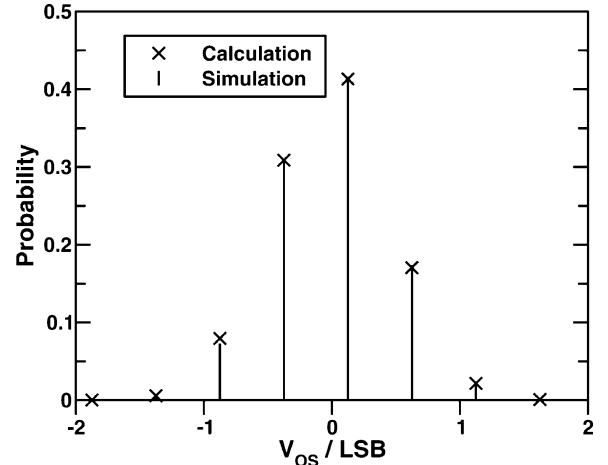


Fig. 6. $M(V_{OS})$ of a BCC example. $\Delta V = (1/2)$ LSB, $N_C = 64$, and $V_{OS}^0 = (1/4)\Delta V$.

and the smooth dashed line is the approximation using (3). The settling time constant is $\tau_c = \pi \cdot 2^6 \cdot 64 \approx 12868$.

B. Offset Fluctuation

As the V_{OS} is converged toward zero by the calibration process, the behavior of $V_{OS}[k]$ becomes a discrete random fluctuation around zero. The stochastic behavior of V_{OS} is a random sequence $V_{OS}[k]$. Fig. 5 illustrates a possible probability mass function (PMF) for V_{OS} , $M(V_{OS})$. The discrete events for V_{OS} , are $V_{OS}^0, V_{OS}^{-1}, V_{OS}^{-2}, V_{OS}^{-3}, \dots$, with V_{OS}^0 being closest to zero. The distance between two adjacent events is ΔV . The possible value for V_{OS} is between $-\Delta V/2$ and $+\Delta V/2$. The calibration loop forces the maximum value of $M(V_{OS})$ to occur at V_{OS}^0 . A mathematical treatment of $V_{OS}[k]$ is included in the Appendix, which also includes the procedures to calculate $M(V_{OS})$ from ΔV , N_C , and the V_i 's PDF.

Fig. 6 shows the $M(V_{OS})$ of a BCC with condition identical to the one for Fig. 4. Results from both calculation and simulation are presented. The value of V_{OS}^0 is chosen to be $(1/4)\Delta V$. As expected, the maximum probability for V_{OS} occurs at V_{OS}^0 . The V_{OS} can appear at other values away from zero, but with diminishing probability.

From $M(V_{OS})$, both the mean, $\mu(V_{OS})$, and standard deviation, $\sigma(V_{OS})$, of V_{OS} can be calculated. The value of $\mu(V_{OS})$ is always zero, which is enforced by the calibration feedback mechanism. The value of $\sigma(V_{OS})$ depends on ΔV , N_C , and

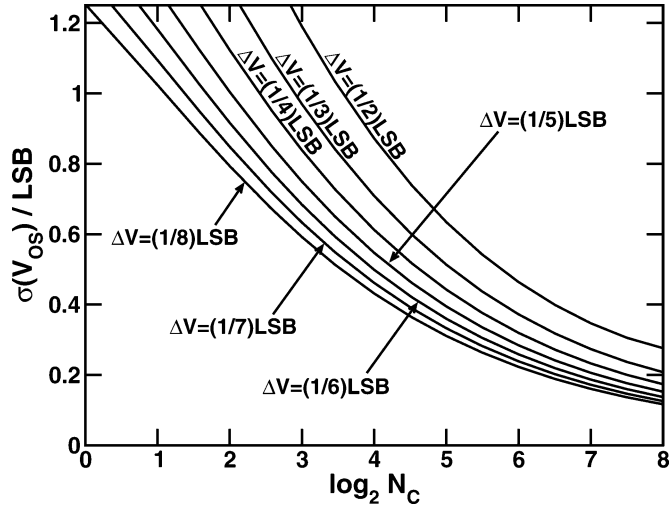


Fig. 7. $\sigma(V_{OS})$ of a BCC example. Assume that the 6-bit ADC design case, but with various values of N_C and ΔV .

V_{OS}^0 . If V_{FS} is much larger than ΔV so that $P_1 \gg \Delta P_1$, the effect of V_{OS}^0 on $\sigma(V_{OS})$ becomes insignificant.

Fig. 7 shows the ΔV and N_C dependence of $\sigma(V_{OS})$ for a middle BCC in the 6-bit ADC design case. The V_{OS}^0 dependence of $\sigma(V_{OS})$ is neglected since $P_1 \gg \Delta P_1$. The $\sigma(V_{OS})$ can be reduced by decreasing ΔV or increasing N_C , but at the expense of increasing time constant τ_c of (3). To achieve $\sigma(V_{OS}) < (1/3)$ LSB in the 6-bit ADC design case, one can choose $\Delta V = (1/8)$ LSB and $N_C = 2^5$ for short τ_c , or $\Delta V = (1/2)$ LSB and $N_C = 2^8$ for long τ_c . At the circuit level, a smaller ΔV is more difficult to implement and more digital bits are required to maintain similar range for offset control.

IV. FLASH ADC USING WINDOWED BCCS

The BCC shown in Fig. 3, which is consisted of an RCC and a CP, is designed to replace every comparator in a flash ADC. In each BCC, its CP is activated if the corresponding output $D_c = 1$, i.e., V_i is in the $(P_1 + \Delta P_1)$ regions of Fig. 2. In a flash ADC, the $V_{R,j}$ is different for different BCC, and the input signal may not be a sine wave in practical applications. Therefore, several issues need to be addressed.

- 1) *Null-information input condition.* When the input V_i does not have value near $V_{R,j}$ for a long period of time, the j th BCC experiences a condition with $\Delta P_1 = 0$. The corresponding CP receives no meaningful information about V_{OS} . But since $P_1 \neq 0$, the V_{OS} of the j -th BCC may wander around the null-information region where $\Delta P_1 = 0$. This phenomenon of V_{OS} wandering does not affect the quality of the ADC's output as long as ΔP_1 remains zero. However, as soon as the input condition is changed so that $\Delta P_1 \neq 0$, the ADC may suffer a large V_{OS} at the j th BCC before its CP can make the necessary correction.
- 2) *Small $\Delta P_1/P_1$ ratio.* The V_{OS} fluctuation is related to the $\Delta P_1/P_1$ ratio. When the ratio is small, most of variation in V_{OS} is of no significance, and just introduces fluctuation.

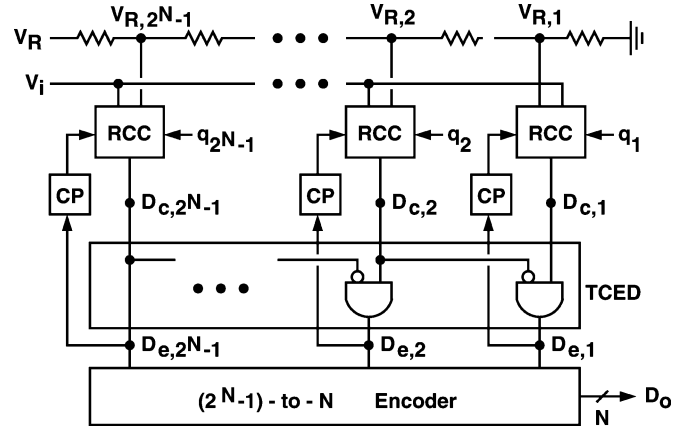


Fig. 8. Flash ADC employing windowed BCCs.

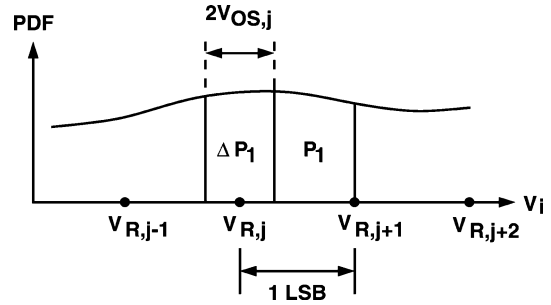


Fig. 9. Windowing effect for j th BCC.

- 3) *P_1 variation.* In a flash ADC, different comparators are associated with different $V_{R,j}$ values, thus perceives drastically different P_1 values even if the input's PDF has an uniform distribution. The P_1 variation leads to different $\Delta P_1/P_1$ ratio perceived by different comparators. This can complicate the design of N_C and ΔV parameters for each BCC when V_{OS} fluctuation is considered.

The above issues can be resolved by rearranging the BCCs as shown in Fig. 8. In this architecture, the outputs from all BCCs, $D_{c,j}$ for $1 \leq j \leq 2^N - 1$, are fed into a thermometer-code edge detector (TCED) to generate an edge code $D_{e,j}$ for $1 \leq j \leq 2^N - 1$. In Fig. 8, the TCED simply consists of two-input AND gates, so that $D_{e,j} = 1$ if $D_{c,j} = 1$ and $D_{c,j+1} = 0$. The CP of the j th BCC uses $D_{e,j}$ as its input, instead of $D_{c,j}$. With this arrangement, the j th CP is activated only when V_i appears between $V_{R,j} + q_j \cdot V_{OS,j}$ and $V_{R,j+1} + q_{j+1} \cdot V_{OS,j+1}$. In most flash ADC designs, the TCED is often a sub-block of the back-end output encoder, thus no extra hardware is required for the proposed architecture.

The arrangement of Fig. 8 introduces a windowing effect which can reduce the P_1 value perceived by each BCC. As illustrated in Fig. 9, the P_1 region for the j th BCC is now confined between $V_{R,j} + V_{OS,j}$ and $V_{R,j+1}$. The actual upper bound for the P_1 region is either $V_{R,j+1} - V_{OS,j+1}$ or $V_{R,j+1} + V_{OS,j+1}$ depending on the q_{j+1} random sequence. The averaged value of $V_{R,j+1}$ is used in Fig. 9. With this windowing mechanism, the $\Delta P_1/P_1$ ratio for every BCC is drastically increased, resulting in smaller V_{OS} fluctuation. In addition, the difference in $\Delta P_1/P_1$ between different BCCs is also reduced unless the

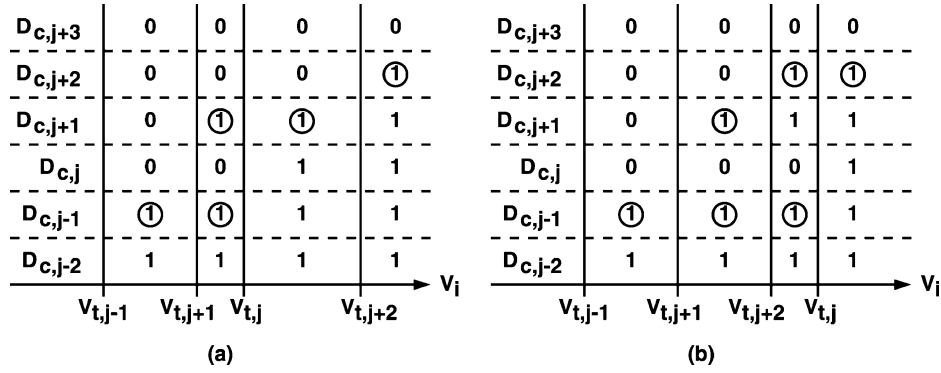


Fig. 10. NMT examples. (a) $V_{t,j+1} < V_{t,j}$. (b) $V_{t,j+2} < V_{t,j}$.

PDF of the ADC's input has a drastically nonuniform distribution. Then, all BCCs in a ADC can employ identical N_C and ΔV for offset calibration.

It is essential that the q_j and q_{j+1} random sequences be mutually uncorrelated. Otherwise, the averaged values of P_1 perceived by the j th BCC will be different for different time period, depending on status of q_{j+1} being +1 or -1. Then the entire calibration process can no longer function properly, since it is based on the assumption of static probabilities.

The phenomenon of V_{OS} wandering due to null-information input condition can still occur. However, if ΔV is assumed to be infinitesimal, the maximum distance $V_{OS,j}$ can wander around is 1 LSB, which is the difference between $V_{R,j}$ and $V_{R,j+1}$. Whenever $V_{OS,j} \geq 1$ LSB, P_1 becomes zero and $V_{OS,j}$ can no longer wander. Thus, for a properly designed ADC of Fig. 8, its worst case differential nonlinearity (DNL) is 1 LSB plus the ΔV step size.

The windowing effect introduced by the TCED becomes complicated when the threshold levels of the BCC array are not monotonic. The threshold of the j th BCC can be expressed as

$$V_{t,j} = V_{R,j} + q_j \times V_{OS,j} \quad (4)$$

where $q_j \in \{+1, -1\}$ is a random sequence. The difference between two adjacent reference levels is 1 LSB, i.e., $V_{R,j+1} - V_{R,j} = 1$ LSB. Under normal conditions with $V_{OS,j} \ll 1$ LSB, we have $V_{t,j-1} < V_{t,j} < V_{t,j+1}$ for all j . However, during the initial phase of calibration, it is possible that $V_{OS,j} > (1/2)$ LSB at some locations. Then, the non-monotonic-threshold (NMT) condition may occur, in which $V_{t,j} > V_{t,j+m}$ with $m > 0$ for some specific j values.

Fig. 10 shows two examples of NMT conditions. The x axis is the V_i input, tagged with threshold levels of the BCC array. The binary numbers above the x axis indicate the digital outputs of the BCC array. An encircled "1" indicates that the corresponding TCED output is activated. Using Fig. 10(a) as an example, if V_i appears between $V_{t,j-1}$ and $V_{t,j+1}$, we have $D_{c,m} = 1$ for $m \leq j-1$, $D_{c,m} = 0$ for $m \geq j$, and only $D_{c,j-1}$ of all TCED outputs has a value of "1". As a result, only the $(j-1)$ th CP can get activated. If V_i is between $V_{t,j+1}$ and $V_{t,j}$, we have both $D_{c,j-1} = 1$ and $D_{c,j+1} = 1$, due to the NMT condition with $V_{t,j+1} < V_{t,j}$. Thus, both the $(j-1)$ th and $(j+1)$ th CPs are activated simultaneously. For the $(j-1)$ th and $(j+1)$ th CPs, the V_i range in which they are activated is widened, resulting in a larger P_1 probability value as illustrated in Fig. 9.

It can also be found from Fig. 10(a) that the j th CP never gets activated. For the Fig. 10(b) case, in which $V_{t,j+2} < V_{t,j}$, the V_i activation region for the $(j-1)$ th CP is widened further, and the j th CP also never gets activated.

Therefore, with a windowing TCED, the NMT condition can increase the P_1 value for some BCCs, and stop the calibration process for other BCCs. Larger P_1 value perceived by a BCC causes larger V_{OS} fluctuation. But the above condition occurs only during the initial phase of calibration. Once the values of $|V_{OS,j}|$ for all j are trimmed below $(1/2)$ LSB, the monotonicity of threshold levels can be ensured, and the P_1 value perceived by each BCC is reduced to the simple case as shown in Fig. 9.

On the other hand, if some of CPs are always disabled by the NMT effect, the corresponding BCCs then never have the chance to adjust their offsets. It is possible that the NMT condition that causes this inactivity also remains unchanged. To prevent this deadlock, the random sequences, as expressed by the q_j term in (4), need to be uncorrelated among the neighboring BCCs. The use of uncorrelated random sequences can scramble the relative $V_{t,j}$ positions of the neighboring BCCs so that no CP is disabled permanently. Although $2^N - 1$ random sequences q_j for $1 \leq j \leq 2^N - 1$ are shown in Fig. 8, it is sufficient to break the deadlock by letting the random sequences to be uncorrelated only between the adjacent BCCs. Thus, one can choose $q_1 = q_3 = q_5 = \dots$ and $q_2 = q_4 = q_6 = \dots$.

A. Transient Behavior

From (2), the BCC's transient behavior is a function of ΔP_1 , and is not affected by any change in P_1 . Thus, (3) can still be used to estimate the settling time of the calibration process for the BCCs in Fig. 8. The τ_c time constant can be different for different BCC if the V_i input has a nonuniform PDF. A simple approximation is by assuming V_i is uniformly distributed over the entire input range, so that $D(V_{R,j}) = 1/V_{FS}$. Then, the time constant for all BCCs can be expressed as

$$\tau_c = N_C \times \frac{V_{FS}}{\Delta V}. \quad (5)$$

The actual transient response can be slower than the one predicted by (5) if the NMT condition occurs. Since some of the CPs are disabled by the NMT effect, it takes longer time for the entire BCC array to converge.

All BCCs in Fig. 8 can be calibrated independently and simultaneously, thus the settling time is independent of the number of

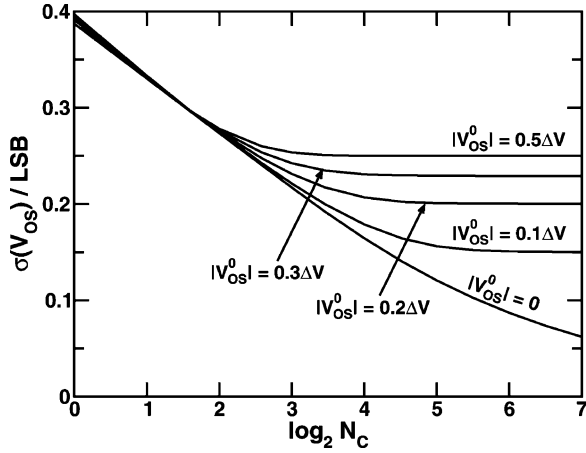


Fig. 11. $\sigma(V_{OS})$ of a windowed BCC example. Assume a 6-bit ADC using windowed BCCs. $\Delta V = (1/2)$ LSB.

BCCs. It is a good practice to install additional power-on calibration mechanism so that the offsets of all comparators are minimized immediately after the hardware power-on phase and before applying the input.

B. Offset Fluctuation

It is difficult to analyze the V_{OS} fluctuation behavior under the NMT condition because neighboring BCCs can interfere with each other. Since $V_{OS,j}$ for all j is to be trimmed to less than $(1/2)$ LSB after the calibration process has converged, the following fluctuation analysis neglects the NMT effect.

With the above assumptions, the fluctuation analysis for the j th BCC is reduced to the calculation of the probability mass function $M(V_{OS})$ for a simple BCC. From Fig. 9, we have $\Delta P_1/P_1 = 2V_{OS}/(1 \text{ LSB} - V_{OS})$. It is assumed that the input's PDF is uniform within the window. Then, we can calculate $M(V_{OS})$ using the procedures described in the Appendix. Notably, the $M(V_{OS})$ depends on $\Delta P_1/P_1$, ΔV , and N_C only. The larger probability of $U = 0$ introduced by the windowing effect does not affect $M(V_{OS})$.

Fig. 11 shows the calculated V_{OS} standard deviation, $\sigma(V_{OS})$, of a windowed BCC. The 6-bit ADC design case and $\Delta V = (1/2)$ LSB are assumed. Unlike the calculation results shown in Fig. 7, the influence of the minimum offset, V_{OS}^0 , on $M(V_{OS})$ becomes evident due to a much larger value of the $\Delta P_1/P_1$ ratio. As N_C increases, the $\sigma(V_{OS})$ decreases and is saturated at different value for different V_{OS}^0 . The worst case saturation value for the $\sigma(V_{OS})$ occurs at $|V_{OS}^0| = (1/2)\Delta V$. In this case and with large N_C , the V_{OS} stays at either $+(1/2)\Delta V$ or $-(1/2)\Delta V$, resulting in $\sigma(V_{OS}) = (1/2)\Delta V = (1/4)$ LSB.

Fig. 12 shows the ΔV and N_C dependence of the $\sigma(V_{OS})$. The 6-bit ADC case using windowed BCCs and the worst case $|V_{OS}^0| = (1/2)\Delta V$ are assumed. Comparing with Fig. 7, the $\sigma(V_{OS})$ is reduced drastically by using windowed BCCs with identical ΔV and N_C .

C. 6-bit ADC Design Case

The design example is a 6-bit ADC based on Fig. 8. All BCCs are identical and have $\Delta V = (1/4)$ LSB and $N_C = 16$. The design also includes a thermometer-to-Gray encoder and a

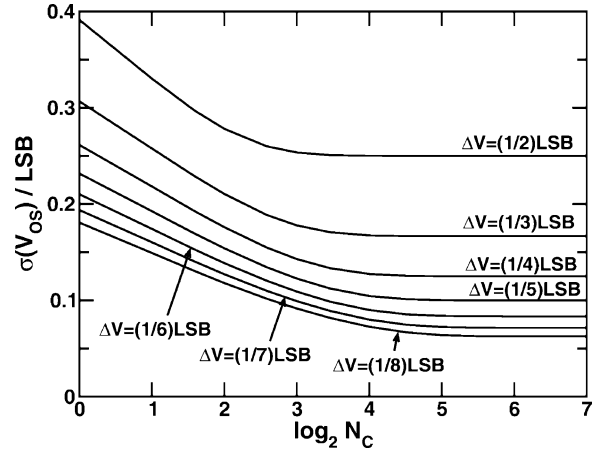


Fig. 12. Worst case $\sigma(V_{OS})$ for various values of ΔV and N_C . Assume a 6-bit ADC using windowed BCCs and $V_{OS}^0 = (1/2)\Delta V$.

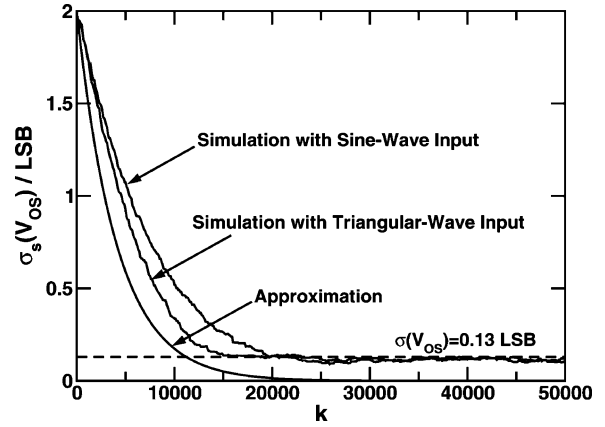


Fig. 13. Transient behavior of a 6-bit ADC design case.

Gray-to-binary encoder to mitigate the bubble-error effect. All BCCs are introduced with initial offsets, which have a random Gaussian distribution with a standard deviation of 2 LSB. The ADC's input is a full-range sine wave in the simulations.

Fig. 13 shows the transient behavior of the V_{OS} spatial standard deviation, $\sigma_s(V_{OS})$, from the simulation of the 6-bit ADC example. The spatial standard deviation is collected by recording V_{OS} of all BCCs at a given time. It is a good approximation for individual BCCs $\sigma(V_{OS})$, since the random variable V_{OS} is independent and ergodic. In Fig. 13, the $\sigma_s(V_{OS})$ is initially set at 2 LSB, then as calibration proceeds, is forced to settle to a nonzero value close to 0.13 LSB. This steady-state value, as indicated by the horizontal dashed line, is obtained from Fig. 12. As expected, the ADC's input can affect the calibration transient behavior. Simulation with triangular-wave input, which has an uniform PDF, settles faster than the case with sine-wave input. Also plotted in Fig. 13 is the approximation of (5) and (3). The time constant is $\tau_c = N_C \times 2^6 \times 4 = 4096$. From (3), it will take $2.73\tau_c$, which is approximately 11 000 samples, for the ADC to reduce its $\sigma_s(V_{OS})$ from 2 LSB to 0.13 LSB. On the other hand, the corresponding settling time is approximately 15,000 samples in simulation with triangular-wave input and 20 000 samples in simulation with sine-wave input. The settling time deviation of the triangular-wave case is due to the NMT effect, which

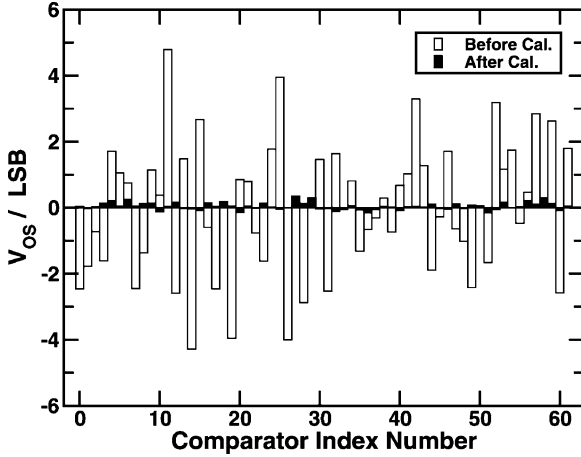


Fig. 14. Offsets of windowed BCCs before and after calibration.

temporarily inhibits some of the BCCs from self-adjusting, thus slows down the entire calibration process.

The effectiveness of the proposed calibration scheme is demonstrated in Fig. 14. Shown in the figure are offsets of each BCC before and after calibration. Data are recorded at $k = 0$ and $k = 1\,000\,000$ respectively. Offsets on all comparators are trimmed to a value below 0.5 LSB.

V. CONCLUSION

A background calibration technique is described for trimming the input-referred offsets of the comparators in a flash ADC. For a RCC, the polarity of its offset is detected by observing the code density of its comparison results. Binary feedback is then used to digitally adjust the comparator's offset so that the offset is minimized. All calibration procedures are performed in the digital domain.

There are two key design parameters, i.e., the step size of the comparator's offset control, ΔV , and the threshold of the bilateral peak detector, N_C . For the j th comparator, the time constant of its calibration loop is proportional to the $N_C/\Delta V$ ratio and is inversely proportional to the probability for the ADC's input, V_i , appearing near the j th reference level, $V_{R,j}$. Once the calibration has converged, the offset of a comparator begins to fluctuate around the zero value due to the disturbance of V_i . The standard deviation of this fluctuation, $\sigma(V_{OS})$, can be calculated from N_C , ΔV and the V_i 's PDF. The $\sigma(V_{OS})$ can be reduced by increasing N_C or decreasing ΔV .

The $\sigma(V_{OS})$ of a BCC can be drastically reduced by selecting only relevant comparison results for calibration input. This can be accomplished by observing the entire comparator array's outputs to determine if V_i appears near the $V_{R,j}$ reference level for the j th BCC. The use of the above windowing scheme to reduce $\sigma(V_{OS})$ may disable some BCC to perform calibration due to the NMT condition. The use of uncorrelated random chopping for neighboring BCCs can ensure the validity of offset detection, and mitigate the NMT effects, which include an increase of calibration settling time and an increase of $\sigma(V_{OS})$.

For a 6-bit flash ADC design example, $\Delta V = (1/4)$ LSB and $N_C = 16$ are chosen to calibrate all 63 windowed BCCs. When a full-range sinusoidal input is applied, offsets of all BCCs are

reduced automatically from $\sigma(V_{OS}) = 2$ LSB to $\sigma(V_{OS}) = 0.13$ LSB in a time period of 20 000 samples.

The proposed calibration scheme uses extra digital hardware to overcome the inherent speed-power-accuracy limitation of MOS transistors in comparator designs. The scheme is most suitable for use in advanced CMOS technologies, in which device scaling diminishes the cost of digital circuits.

APPENDIX

The appendix describes a mathematical treatment of the stochastic behavior of V_{OS} in a single BCC. As the calibration converges, V_{OS} varies over a set of discrete values and forms a random variable \mathbf{V}_{OS} , which can be expanded as a discrete random sequence $\mathbf{V}_{OS}[k]$ along the sampling instant on time axis, k . $\mathbf{V}_{OS}[k]$ is a function of three different random sequences, $\mathbf{U}[k]$, $\mathbf{R}[k]$ and $\mathbf{S}[k]$, which stand for the possible values of U , R and S in Fig. 3. The relation between $\mathbf{V}_{OS}[k]$ and $\mathbf{S}[k]$ can be expressed as:

$$\mathbf{V}_{OS}[k] = \mathbf{V}_{OS}[k-1] + \mathbf{S}[k] \times \Delta V \quad (6)$$

where $\mathbf{S}[k] \in \{-1, 0, +1\}$ and is determined by the probability mass function (PMF) of $\mathbf{R}[k]$, $M(\mathbf{R}[k])$. On the other hand, $\mathbf{U}[k] \in \{-1, 0, +1\}$ is the result of a multinomial trial. Its PMF can be calculated from P_1 and ΔP_1 that describe the V_i input probabilistic distribution as illustrated in Fig. 2. We have

$$\begin{aligned} M(\mathbf{U}[k] = +1) &= \frac{1}{2}P_1 \\ M(\mathbf{U}[k] = 0) &= 1 - P_1 - \frac{1}{2}\Delta P_1 \\ M(\mathbf{U}[k] = -1) &= \frac{1}{2}P_1 + \frac{1}{2}\Delta P_1. \end{aligned} \quad (7)$$

Since $\mathbf{U}[k] = 0$ does not effect the value of \mathbf{R} , The value of $M(\mathbf{U}[k] = 0)$ has no influence on $M(\mathbf{V}_{OS})$. Only the conditional probability $M_c(\mathbf{U}[k])$, which stands for the condition $\mathbf{U}[k] = \pm 1$, need to be considered. From (7), $M_c(\mathbf{U}[k])$ can be expressed as

$$\begin{aligned} M_c(\mathbf{U}[k] = -1) &= \frac{P_1}{2P_1 + \Delta P_1} = \frac{1}{2 + \frac{\Delta P_1}{P_1}} \\ M_c(\mathbf{U}[k] = +1) &= \frac{P_1 + \Delta P_1}{2P_1 + \Delta P_1} = \frac{1 + \frac{\Delta P_1}{P_1}}{2 + \frac{\Delta P_1}{P_1}}. \end{aligned} \quad (8)$$

Notably, the $M_c(\mathbf{U}[k])$ is a function of the $\Delta P_1/P_1$ ratio only, and the P_1 and ΔP_1 probabilities have V_{OS} dependence.

Fig. 15 shows the dynamics of $V_{OS}[k]$ probabilistic behavior. Under the condition in which $V_{OS}[k] = V_{OS}^m$, the conditional probability for $S = -1$ is defined as $P_d^m[k]$, the conditional probability for $S = 0$ is defined as $P_s^m[k]$, and the conditional probability for $S = +1$ is defined as $P_u^m[k]$. They are called conditional transfer probabilities. When the calibration converges, $M(\mathbf{V}_{OS}[k])$ becomes time invariant, and the following steady-state properties are assumed.

- 1) *Transfer probability equality*: The probability for V_{OS} moving from V_{OS}^m to V_{OS}^{m+1} is the same as that moving from V_{OS}^{m+1} to V_{OS}^m .

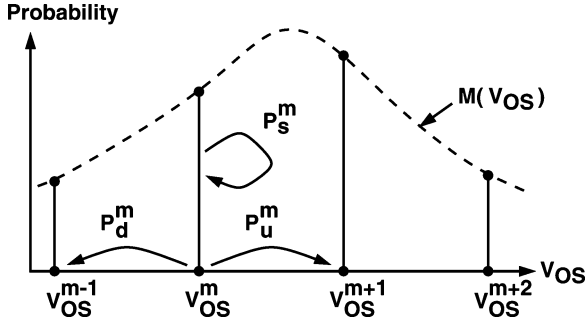


Fig. 15. Steady-state transfer probabilities at V_{OS}^m .

- 2) *Probability conservation*: The probability for V_{OS} moving out of V_{OS}^m is the same as that moving into V_{OS}^m .
- 3) *Probability invariance*: Probabilities in the above properties remain constant over k .

The third property implies that the conditional transfer probabilities $P_d^m[k]$, $P_s^m[k]$, and $P_u^m[k]$ are also constant over k .

Observing that the behavior of $V_{OS}[k]$ depends on the value of R , and even the same R value may come from different historical accumulation of U , we decompose the transfer probabilities into historically unified fractional terms, $\hat{P}_s^m[i]$, $\hat{P}_u^m[i]$, and $\hat{P}_d^m[i]$, which account for the same move-in-and-stay events for V_{OS} . The historically unified fractional terms are defined as follows: after a time period of i cycles for which V_{OS} remains at V_{OS}^m , i.e., $S = 0$, $\hat{P}_s^m[i]$ is the probability of $S = 0$ at cycle $i + 1$, $\hat{P}_u^m[i]$ is the probability of $S = +1$ at cycle $i + 1$, and $\hat{P}_d^m[i]$ is the probability of $S = -1$ at cycle $i + 1$. Note that the definitions exclude the probability of V_{OS} moving into V_{OS}^m from different values in this time period of i cycles. The transfer probabilities illustrated in Fig. 15 can then be expressed as:

$$\begin{aligned} P_s^m &= P^m \sum_{i=0}^{\infty} \hat{P}_s^m[i] \\ P_u^m &= P^m \sum_{i=0}^{\infty} \hat{P}_u^m[i] \\ P_d^m &= P^m \sum_{i=0}^{\infty} \hat{P}_d^m[i] \end{aligned} \quad (9)$$

where P^m is a constant probability for which V_{OS} is adjusted and becomes V_{OS}^m .

The $\hat{P}^m[i]$ terms in the right-hand side of (9) can be calculated by considering all possible conditions for R , and are expressed as:

$$\begin{aligned} \hat{P}_s^m[i] &= \sum_{-N_C \leq R[i] \leq N_C} \hat{M}^m(\mathbf{R}[i]) \\ \hat{P}_u^m[i] &= \hat{M}^m(\mathbf{R}[i]) \Big|_{R[i]=+(N_C+1)} \\ \hat{P}_d^m[i] &= \hat{M}^m(\mathbf{R}[i]) \Big|_{R[i]=-(N_C+1)} \end{aligned} \quad (10)$$

where $\mathbf{R}[i]$ is the random variable of R . $\hat{M}^m(\mathbf{R}[i])$ is the PMF of $\mathbf{R}[i]$ after $V_{OS} = V_{OS}^m$ for consecutive i cycles. $\hat{M}^m(\mathbf{R}[i])$ can be calculated by using the following recursive function [14]:

$$\hat{M}^m(\mathbf{R}[i]) = \left[\hat{M}^m(\mathbf{R}[i-1]) \times W(\mathbf{R}) \right] * M_c^m(\mathbf{U}) \quad (11)$$

where $M_c^m(\mathbf{U})$ is defined in (8) with superscript m denoting $V_{OS} = V_{OS}^m$. Note that $\hat{M}^m(\mathbf{R}[0]) = M_c^m(\mathbf{U})$ for $i = 0$. The operator “*” is the probability convolution over the random variable. $W(\mathbf{R})$ is a window function which has the value 1 if $|R| \leq N_C$ and 0 otherwise. Multiplication by $W(\mathbf{R})$ is to exclude the probability of V_{OS} moving out of V_{OS}^m .

From (9), (10), and (11), we can calculate the transfer probabilities P_s^m , P_u^m , and P_d^m . They are all proportional to P^m . The value of P^m can be found by letting the summation of P_s^m , P_u^m , and P_d^m be equal to one. Although (9) involves summation of infinite number of terms, the values of $\hat{P}_s^m[i]$, $\hat{P}_u^m[i]$, and $\hat{P}_d^m[i]$ approach zero for large i . Thus, the transfer probabilities, P_s^m , P_u^m and P_d^m , can be calculated to any specific precision with finite number of terms.

According to the steady-state property 2, the probability of V_{OS} changing from V_{OS}^m to V_{OS}^{m+1} is the same as the probability of V_{OS} changing from V_{OS}^{m+1} to V_{OS}^m , i.e.,

$$M(V_{OS} = V_{OS}^m) \times P_u^m = M(V_{OS} = V_{OS}^{m+1}) \times P_d^{m+1}. \quad (12)$$

Thus, we have

$$\frac{M(V_{OS} = V_{OS}^m)}{M(V_{OS} = V_{OS}^{m+1})} = \frac{P_d^{m+1}}{P_u^m}. \quad (13)$$

The above equation can be used to calculate $M(V_{OS})$. The absolute values of $M(V_{OS})$ can be found by the fact that integration of $M(V_{OS})$ over all V_{OS} is equal to one.

An example of calculated $M(V_{OS})$ is illustrated in Fig. 6. Also shown in Fig. 6 is the results from simulation using identical design parameters. The simulation result is the normalized histogram of 10^8 samples. Good agreement is demonstrated between the calculation and simulation.

REFERENCES

- [1] K. Bult and A. Buchwald, “An embedded 240-mW 10-b 50-MS/s CMOS ADC in 1-mm²,” *IEEE J. Solid-State Circuits*, vol. 32, no. 12, pp. 1887–1895, Dec. 1997.
- [2] B. P. Brandt and J. Lutsky, “A 75-mW, 10-b, 20-MSPS CMOS sub-ranging ADC with 9.5 effective bits at Nyquist,” *IEEE J. Solid-State Circuits*, vol. 34, no. 12, pp. 1788–1795, Dec. 1999.
- [3] K. Uyttenhove and M. S. J. Steyaert, “Speed-power-accuracy tradeoff in high-speed CMOS ADCs,” *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 49, no. 4, pp. 280–286, Apr. 2002.
- [4] B. Razavi, *Principles of Data Conversion System Design*. New York: IEEE Press, 1995.
- [5] H. Pan and A. A. Abidi, “Spatial filtering in flash A/D converters,” *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 50, no. 8, pp. 424–436, Aug. 2003.
- [6] H. Okada, Y. Hashimoto, K. Sakata, T. Tsukada, and K. Ishibashi, “Offset calibrating comparator array for 1.2-V, 6-bit, 4-Gsample/s flash ADCs using 0.13- μ m CMOS technology,” in *Proc. ESSCIRC’03*, Sep. 2003, pp. 711–714.
- [7] Y. Tamba and K. Yamakido, “A CMOS 6b 500 MSample/s ADC for hard disk drive read channel,” in *Proc. IEEE Int. Solid-State Circuits Conf.*, Feb. 1999, pp. 324–325.
- [8] M. P. Flynn, C. Donovan, and L. Sattler, “Digital calibration incorporating redundancy of flash ADCs,” *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 50, no. 5, pp. 205–213, May 2003.
- [9] M.-J. Choe, B.-S. Song, and K. Bacrania, “A 13-b 40-Msample/s CMOS pipelined folding ADC with background offset trimming,” *IEEE J. Solid-State Circuits*, vol. 35, no. 12, pp. 1781–1789, Dec. 2000.
- [10] H. van der Ploeg, G. Hoogzaad, H. A. H. Termeer, M. Vertregt, and R. L. J. Roovers, “A 2.5v 12-b 54-Msample/s 0.25- μ m CMOS ADC in 1-mm² with mixed-signal chopping and calibration,” *IEEE J. Solid-State Circuits*, vol. 36, no. 12, pp. 1859–1867, Dec. 2001.

- [11] S. M. Jamal, D. Fu, N. C.-J. Chang, P. J. Hurst, and S. H. Lewis, "A 10-b 120-Msample/s time-interleaved analog-to-digital converter with digital background calibration," *IEEE J. Solid-State Circuits*, vol. 37, no. 12, pp. 1618–1627, Dec. 2002.
- [12] M. Q. Le, P. J. Hurst, and K. C. Dyer, "An analog DFE for disk drives using a mixed-signal integrator," in *Proc. Symp. VLSI*, Jun. 1998, pp. 156–157.
- [13] K. Dyer, D. Fu, S. Lewis, and P. Hurst, "Analog background calibration of a 10b 40 MSample/s parallel pipelined adc," in *Proc. ISSCC*, Feb. 1998, pp. 142–427.
- [14] H. Stark and J. W. Woods, *Probability and Random Processes With Application to Signal Processing*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 2002.



Chun-Cheng Huang (S'02) was born in Chia-Yi, Taiwan, R.O.C., in 1970. He received the B.S. degree in electrophysics from National Chiao-Tung University, Hsin-Chu, Taiwan, R.O.C., in 1992, and the M.S. degree in electrical engineering from National Don Hwa University, Hualien, Taiwan, R.O.C., in 1999, respectively. He is currently working toward the Ph.D. degree in the field of high-speed data conversion circuits at National Chiao Tung University.

From 1992 to 1994, he served as an Ordnance Officer in the R.O.C. Army. Since 1994, he has worked

in the area of analog circuit design.



Jieh-Tsornng Wu (S'83–M'87) was born in Taipei, Taiwan, R.O.C., in 1958. He received the B.S. degree in electronics engineering from National Chiao-Tung University, Hsin-Chu, Taiwan, R.O.C, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1980, 1983, and 1988, respectively.

From 1980 to 1982, he served in the Chinese Army as a Radar Technical Officer. From 1982 to 1988, at Stanford University, he focused his research on high-speed analog-to-digital conversion in CMOS very large-scale integration. From 1988 to 1992, he was a Member of Technical Staff at Hewlett-Packard Microwave Semiconductor Division, San Jose, CA, and was responsible for several linear and digital gigahertz integrated circuit designs. Since 1992, he has been with the Department of Electronics Engineering, National Chiao-Tung University, where he is now a Professor. His current research interests are high-performance mixed-signal integrated circuits.

Dr. Wu is a member of Phi Tau Phi.