Design Considerations for Interleaved ADCs

Behzad Razavi, *Fellow, IEEE*

***Abstract*—Interleaving can relax the power-speed tradeoffs of analog-to-digital converters and reduce their metastability error rate while increasing the input capacitance. This paper quantifies the benefits and derives an upper bound on the performance by considering kT/C noise and slewing requirements of the circuit driving the system. A frequency-domain analysis of interleaved converters is also presented that sheds light on the corruption mechanisms due to interchannel mismatches. A background timing mismatch calibration technique is proposed and experimentally shown to reduce the image to -75 dB for input frequencies exceeding 500 MHz.**

***Index Terms*—Background calibration, flash analog-to-digital converters (ADCs), figure of merit, image, interleaving mismatches, low-power ADCs, pipelined ADCs, SAR ADCs, timing mismatch.**

I. Introduction

THE concept of time interleaving was originally proposed as a means of increasing the speed of analog-to-digital converters (ADCs) [1], albeit with a power and area penalty. However, it has since been recognized that interleaving offers additional benefits even if absolute speed is not of primary interest. Such benefits accrue at the cost of other complexities, such as the need for interchannel mismatch correction [2]–[11] and multiphase clock generation.

This paper presents a number of new insights into the operation and performance of interleaved ADCs. Specifically, we study the power efficiency and metastability properties of such ADCs and, through a frequency-domain analysis, quantify the effect of interchannel mismatches. We also propose a background calibration technique for timing mismatches that performs detection in the digital domain and correction in the analog domain. The reader is referred to the conference paper [12] for a survey of the prior art.

Section II describes the advantages of interleaving, particularly the improvement of the figure of merit (FOM) and the reduction of metastability error rate. Section III derives an upper bound on the performance of A/D interfaces, demonstrating that the circuit driving an ADC may consume more power than the ADC itself. Section IV analyzes the effect of interchannel mismatches and quantifies the effect of timing mismatch for a random input signal. Section V presents the new timing mismatch calibration technique and Section VI the experimental results for a two-channel 10-bit prototype.

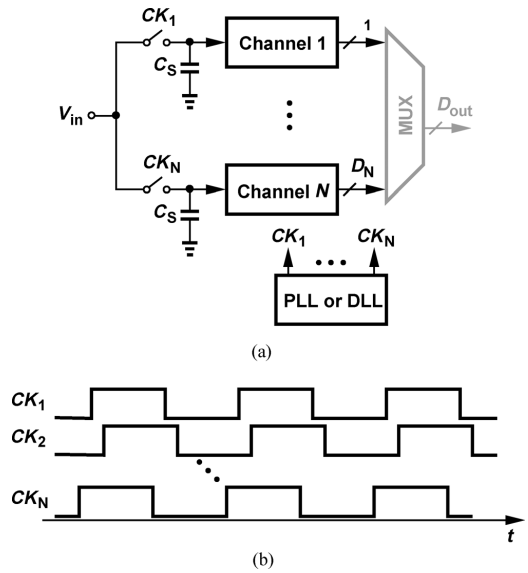


Fig. 1. (a) Interleaving environment, and (b) its clock phases.

II. Advantages of Interleaving

In order to allow greater acquisition and conversion times, N identical ADCs, each having a front-end sampler can be interleaved in the time domain [Fig. 1(a)]. This arrangement requires N clock phases that uniformly span 360° [Fig. 1(b)] and are generated by a phase-locked loop (PLL) or delay-locked loop (DLL). The digital outputs can be multiplexed to reconstruct the data, but in many applications the slower outputs are preferred as they ease the subsequent processing.

In this paper, the maximum analog input frequency is denoted by $f_{\rm in}$, the period of each of the clocks in Fig. 1(b) by $T_{\rm CK} = 1/f_{\rm CK}$, and the overall sampling rate by $f_S = Nf_{\rm CK}$. We assume that the sampling capacitor in each channel is determined by kT/C noise and consider resolutions around 10 bits and sampling rates around 1 GHz. Most of the concepts are presented for $N = 2$ but can be generalized for higher values of N as well.

A. Maximum Speed

The extent to which interleaving improves the conversion rate depends on the relative speeds of the sampler and the quantizer in each channel. In a typical case, the quantizer's long conversion time calls for interleaving, but the maximum number of channels is ultimately limited by the sampler's performance. To

Manuscript received December 03, 2012; revised January 28, 2013; accepted March 20, 2013. Date of publication May 13, 2013; date of current version July 19, 2013. This paper was approved by Guest Editor Ken Suyama.

The author is with the Electrical Engineering Department, University of California, Los Angeles, CA 90095-1594 USA (e-mail: razavi@ee.ucla.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSSC.2013.2258814

formulate this bound, first suppose a single Nyquist-rate channel is designed for *minimum* acquisition time through the use of known techniques such as switch bootstrapping. If the ADC allots $T_{\rm CK}/2 = 1/(2f_S) = 1/(4f_{\rm in})$ seconds to acquisition and requires K time constants, $K\tau_{\rm acq}$, for the resolution of interest, then $T_{\rm CK}/2 = K\tau_{\rm acq}$ and hence

$$f_{\rm in} \leq \frac{1}{4K\tau_{\rm acq}}. \tag{1}$$

In addition, the *small-signal* bandwidth of the sampler (in the acquisition mode) must exceed $f_{\rm in}$ so as to avoid significant attenuation

$$f_{\rm in} \leq \frac{1}{2\pi\tau_{\rm acq}}. \tag{2}$$

For resolutions higher than a few bits, (1) guarantees (2). However, as the number of interleaved channels and hence the maximum input frequency increase, (2) eventually dominates. For N channels, we express (1) as $N/(4K\tau_{\rm acq})$ and equate the result to (2), obtaining

$$N \leq \frac{2K}{\pi}. \tag{3}$$

Thus, *if* the samplers are the speed bottleneck, *then* interleaving beyond $N = 2K/\pi$ provides little improvement. In general, for an M-bit system to settle to 0.5 LSB, we have $K = (M+1)\ln 2$ and hence

$$N < (M+1)\ln 2 \tag{4}$$

$$< 0.44(M+1). \tag{5}$$

For example, a 10-bit ADC designed for maximum speed negligibly benefits from interleaving if $N > 4$.

B. Power-Speed Tradeoffs

An important attribute of time interleaving that was evidently not appreciated even for two decades after Black and Hodges' paper is the flexibility that it affords in the power-speed tradeoff. The work in [13] was perhaps the first to show that the overall power consumption of a particular design reached a *minimum* for four pipelined channels. Recent work on successive-approximation (SAR) ADCs has extensively exploited this attribute to achieve low figures of merit [11], [14]. We define FOM as the power consumption divided by the product of $2^{\rm ENOB}$ and $\min(f_S,\ 2f_{\rm in})$, where ENOB denotes the effective number of bits at an input frequency of $f_{\rm in}$.

Interleaving improves the FOM because, as the conversion speed of a single channel approaches the limits of the technology, the power-speed tradeoff becomes nonlinear, demanding a disproportionately higher power for a desired increase in speed. For example, op amps and comparators eventually reach diminishing returns in their speed as their power consumption is raised. From another perspective, each ADC architecture incurs a certain "timing overhead" that does not easily scale with power.

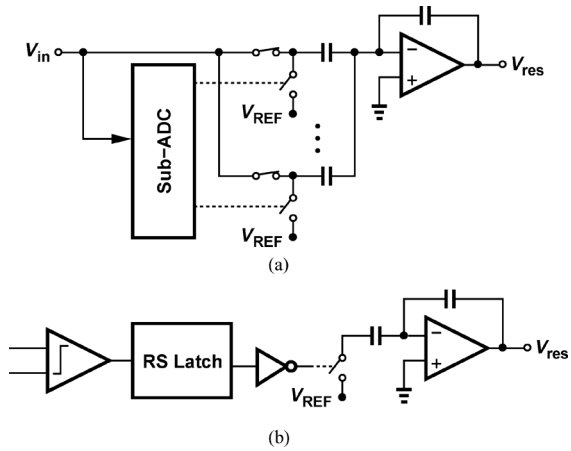


Fig. 2. (a) Pipelined ADC front end, and (b) one slice of sub-ADC driving a unit capacitor in the MDAC.

As an example, consider the first stage of a pipelined ADC, shown in Fig. 2(a), where a sub-ADC resolves a few bits and a multiplying digital-to-analog converter (MDAC) produces the residue, $V_{\rm res}$. A typical design also employs a nonoverlapping clock generator to avoid input-dependent charge injection and ensure that the op amp output does not travel toward the supply rails during clock transitions or the input is not temporarily shorted to the reference(s).

We can identify several bottlenecks that present severe power-speed tradeoffs in the above ADC. 1) The response of the sub-ADC, specifically, that of the comparators, the RS latches, and the buffers necessary to drive the unit capacitors in the MDAC [Fig. 2(b)]; the total delay associated with this path only weakly scales with power. For example, even if a low-power StrongArm comparator [15], [29] (with rail-to-rail outputs) is replaced with a fast, power-hungry current-steering topology [16], the CMOS level conversion that must follow introduces substantial delay. 2) The DAC settling; in addition to the inherent time constant of the unit capacitor paths, the finite output impedance associated with the reference, $V_{\rm REF}$, considerably increases the settling time. The generation of this reference has proved quite challenging [17], [18], especially with low power consumption. 3) The nonoverlap time between the clock phases; due to the length and complexity of the interconnects in the clock distribution network, this "dead" time is difficult to scale down. 4) The finite rise and fall times of the clock; owing to the capacitance of the interconnects and clocked MOSFETs as well as the finite supply bond-wire inductances, transition times typically reach several gate delays, even if large, power-hungry clock buffers are used. As an example, Table I summarizes these timing overhead components for the single-channel 65-nm 10-bit 1-GHz ADC reported in [20]. Not observable externally and hence based on simulations, these values nonetheless point to the difficulty of reducing the timing overhead in practical designs. We remark that each conversion cycle must accommodate 2×20 ps for the nonoverlap time and 2×30 ps for the clock rise and fall times.

As another example, let us consider a generic charge-redistribution SAR ADC (see Fig. 3), whose cycle time is determined

TABLE I
TIMING OVERHEADS FOR THE ADC IN [20]

First Sub−ADC Response Time	150 ps
DAC Time Constant	30 ps
Non−Overlap Time	20 ps
Clock Rise and Fall Times	30 ps

by the comparator delay, the SAR logic delay, and the DAC settling. As with the pipelined ADC studied above, the comparator and the DAC impose a certain weakly-scalable timing overhead. Moreover, the logic is driven by *one* comparator while driving about *half* of the entire DAC capacitance, $C_{\rm DAC}$, in the first conversion step. That is, the logic can be viewed as a single chain that must minimally load the comparator and still drive $C_{\rm DAC}/2$ with a small delay. If tapered uniformly, the logic exhibits a total delay given by $eT_d \ln(0.5C_{\rm DAC}/C_u)$ [21], where T_d is the logic delay with unity fanout and C_u is a capacitance representing the drive capability of the comparator as if it were an inverter driving the chain.[1] This lower bound on the delay in a SAR loop is primarily a function of the technology's gate delay and kT/C noise. For example, if kT/C noise demands a minimum $C_{\rm DAC}$ of 0.5 pF, and if $C_u = 10$ fF, then the logic delay reaches about $9T_d$. Of course, the above three components are multiplied by the number of SAR cycles necessary for one conversion. With 2 bits resolved per step [22], a 10-bit SAR ADC would require 6 steps, incurring a total weakly-scalable logic delay of $54T_d$ in our example plus additional delays due to the comparator and the DAC.

Due to the relatively unscalable timing overheads in each architecture, the figure of merit of ADCs tends to degrade at higher speeds. For example, the FOM rises from 10 fJ per conversion step for the 8-MHz 12-bit ADC in [24] to 500 fJ per conversion step for the 3-GHz 12-bit ADC in [25]. This trend can be ameliorated through the use of time interleaving. In order to quantify this proposition, we assume a simple case where the critical times in an ADC can be lumped into an unscalable overhead, $T_{\rm ov}$, and a linearly power-scalable remainder, $T_{\rm CK} - T_{\rm ov}$. For a single channel, $T_{\rm CK} = T_S$ and $T_S - T_{\rm ov} \propto 1/P_1$ where P_1 denotes the power drawn by the scalable functions in the ADC. The single channel thus demands a power of

$$P_1 = \frac{\alpha}{T_S - T_{\rm ov}} \tag{6}$$

(plus the power consumed by the unscalable functions) to reach a sampling rate of $1/T_S$, where α is a proportionality factor. For N interleaved channels, on the other hand, the scalable remainder is equal to $T_{\rm CK} - T_{\rm ov} = NT_S - T_{\rm ov}$, yielding an overall power consumption of

$$P_N = \frac{\alpha N}{NT_S - T_{\rm ov}} \tag{7}$$

[1]The drive capability of the comparator and hence C_u can be increased by scaling up all of the comparator devices and its power consumption, but the delay has only a logarithmic dependence on C_u.

for the scalable functions. It follows that:

$$\frac{P_N}{P_1} = \frac{N(T_S - T_{\rm ov})}{NT_S - T_{\rm ov}}. \tag{8}$$

This equation formulates the power advantage of interleaving as N increases. For the single-channel 1-GHz example shown in Table I, $T_{\rm ov} \approx 400$ ps, and hence the power drain can be reduced considerably by interleaving.

C. Clock Network Power Consumption

Interleaving can potentially *lower* the power consumption of the clock network. The system of Fig. 1 requires N clock phases that must be distributed across the ADC chip. Let us first consider a single-channel pipelined ADC whose amplifiers, comparators, and capacitors are designed so as to meet certain thermal and kT/C noise requirements. Such a design presents to the clock a total MOS switch capacitance of $C_{\rm MOS}$ and a total interconnect capacitance of $C_{\rm int}$, drawing a power of at least $f_S(C_{\rm MOS} + C_{\rm int})V_{DD}^2$ in the clock path.

Now, we interleave N channels, each running at $f_{\rm CK} = f_S/N$, and make the following observations: 1) determined by kT/C noise, the capacitors in each channel cannot be reduced; 2) the widths of the clocked MOSFETs can be scaled down by a factor of N if we allow all settling times to increase by the same factor[2]; and 3) the total interconnect length per channel is relatively constant because, while most of the MOSFETs are scaled down, the clock phases must still travel from a central point on the chip to all N channels. Thus, the clock power dissipation is now equal to $(f_S/N)N(C_{\rm MOS}/N + C_{\rm int})V_{DD}^2 = f_S(C_{\rm MOS}/N + C_{\rm int})V_{DD}^2$, somewhat lower than that of the single-channel ADC. As a point of reference, $C_{\rm MOS} = 2.5$ pF and $C_{\rm int} = 0.6$ pF in the single-channel 10-bit design reported in [20], suggesting that interleaving can indeed reduce the clock dissipation component. (In practice, timing overheads do not allow the luxury of scaling up all settling times by N, lessening this power saving.)

D. Metastability

Time interleaving can substantially reduce the probability of metastable states in ADCs. This attribute proves particularly useful in multistep (e.g., pipelined or SAR) architectures as they do not permit comparator pipelining (unless the conversion time is increased).[3]

Suppose a single-channel ADC employs a comparator design having a regeneration time constant of $\tau_{\rm reg}$. The probability of a metastable state is given by $\alpha \exp(-T_1/\tau_{\rm reg})$ [19], where α is a proportionality factor and T_1 is the nominal time allocated for the comparator decision. Thus, of a large number of conversions, L, we expect that $\alpha \exp(-T_1/\tau_{\rm reg})L$ are erroneous. Now, we interleave N such ADCs, assuming that the allowable comparator decision time can be increased to NT_1. Each channel therefore exhibits an error rate equal to $\alpha \exp[-(NT_1)/\tau_{\rm reg}]$,

[2]The input sampling devices are an exception as they must provide sufficient bandwidth. In the design in [20], these account for 3% of the clocked MOS gate capacitance.

[3]Feasible in full-flash ADCs, comparator pipelining by adding clocked latches increases the delay in the critical path of multistep ADCs.

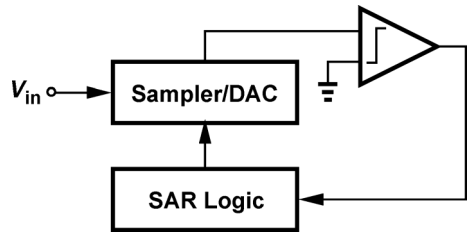


Fig. 3. Generic SAR ADC.

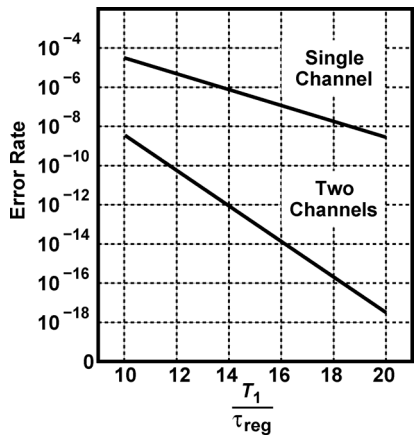


Fig. 4. Reduction of metastability errors as a result of interleaving.

producing $\alpha \exp[-(NT_1)/\tau_{\rm reg}]L$ erroneous outputs for every L conversions.

Upon multiplexing the outputs, we obtain a total of NL conversions, of which $\alpha N \exp[-(NT_1)/\tau_{\rm reg}]L$ are incorrect. That is, the probability of metastable states in the interleaved system is given by

$$P_{{\rm met},N} = \alpha \exp \frac{-NT_1}{\tau_{\rm reg}}. \tag{9}$$

Fig. 4 plots this result as a function of $T_1/\tau_{\rm reg}$ for $N = 1$ and 2, assuming for simplicity that $\alpha = 1$. For example, if $T_1/\tau_{\rm reg} = 12$ (as in the design in [20]), then $P_{{\rm met},N}$ drops from 6×10^{-6} for $N = 1$ to 3.8×10^{-11} for $N = 2$.

The above calculations are not in complete agreement with the power saving derivations in Section II.B: if comparator regeneration is allowed to last NT_1 seconds rather than T_1 seconds, then the time remaining for the other functions is less than that assumed in arriving at (8). However, the dramatic drop suggested by Fig. 4 indicates that the regeneration time need not be increased by a factor of N for acceptable error rates.

III. Disadvantages of Interleaving

In addition to interchannel mismatches and the area penalty, interleaving entails a number of other issues as well.

A. Multi-Phase Clock Generation and Distribution

The large footprint of each channel in an interleaved system inevitably translates to long, complex interconnects for the clock phases (and/or the analog input). Thus, timing mismatches manifest themselves in both the generation and the distribution of the clock phases. For two channels, the use of predictive gating can reduce the timing mismatch to that between only two transistor pairs [29]. For four channels, frequency division provides a compact and efficient solution but for larger N, more complex techniques such as phase interpolation may be necessary [23]. Retiming and gating prove useful in these cases as well [29].

B. Input Capacitance

With kT/C noise imposing a lower bound on each channel's sampling capacitor(s), interleaving proportionately raises the input capacitance. For $N > 2$ channels in Fig. 1, the preceding stage sees a capacitance of $(N/2)C_S$, potentially drawing substantial power. In order to quantify this oft-ignored effect, consider an M-bit ADC sensing an input $V_{\rm in} = V_0 \cos(2\pi f_{\rm in} t)$. The signal-to-noise ratio, SNR, due to quantization and kT/C noise can be expressed as[4]

$$\text{SNR} = \frac{V_0^2/2}{\frac{1}{12}\left(\frac{2V_0}{2^M}\right)^2 + \frac{kT}{C_S}}. \tag{10}$$

Suppose C_S is chosen for an SNR penalty of 0.2 dB, i.e., the second term in the denominator of (10) is 0.047 times the first. It follows that:

$$C_S = \frac{64kT}{V_0^2} 2^{2M}. \tag{11}$$

The stage preceding the ADC must drive this capacitance without slewing (at least at the moment the sampling switch turns off), thereby requiring a bias current given by

$$\frac{I_{SS}}{C_S} \geq V_0(2\pi f_{\rm in}). \tag{12}$$

The power consumption of the driver stage is thus equal to $I_{SS}V_{DD} = V_0 V_{DD}(2\pi f_{\rm in})C_S$ if class-A operation is assumed. As with ADCs, we define a figure of merit, $\text{FOM}_{\rm dr}$, for the driver as power consumption/$(2^M \times f_S)$ and, from (11) and (12), write

$$\text{FOM}_{\rm dr} = \frac{2\pi f_{\rm in} V_0 \frac{64kT}{V_0^2} 2^{2M} V_{DD}}{2^M f_S}. \tag{13}$$

At Nyquist rate, $f_{\rm in} = f_S/2$ and hence

$$\text{FOM}_{\rm dr} = \frac{64\pi kTV_{DD}}{V_0} 2^M. \tag{14}$$

For the driver stage to provide sufficient linearity, its peak-to-peak output swing, $2V_0$, is likely to be no more than $V_{DD}/2$; that is

$$\text{FOM}_{\rm dr} = 256\pi kT \times 2^M. \tag{15}$$

For example, if $M = 12$, $\text{FOM}_{\rm dr} = 13.6$ fJ/conversion step, a value *greater* than the FOM of the 12-bit ADC in [24].

[4]We neglect the kT/C noise due to other capacitors in the signal path.

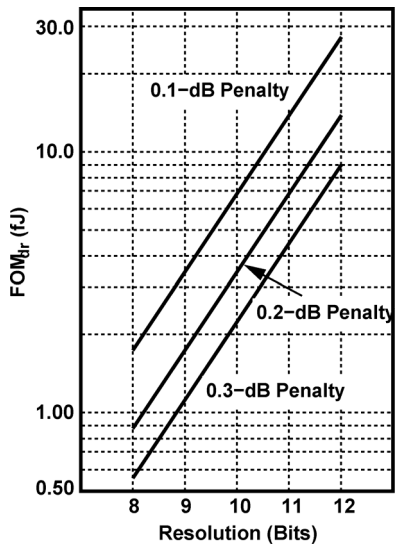


Fig. 5. Driver FOM for different SNR penalties.

Derived from first principles, (15) places an upper bound on the performance of A/D interfaces, indicating that the driver FOM inevitably degrades at higher resolutions. Fig. 5 plots this FOM as a function of the resolution for different SNR penalties. If $N > 2$ channels are interleaved, this FOM is multiplied by $N/2$, further degrading the overall performance.[5] Independent of the conversion rate, $\mathrm{FOM_{dr}}$ suggests that, with the rapid advances in interleaved ADCs, the power consumption of the driver may exceed that of the ADC itself! Note that if higher SNR penalties are allowed, then the required resolution cannot be met. While an SNR penalty of 0.2 or 0.3 dB may appear conservative, we point out that if the penalty is allowed to rise to, say, 3 dB, then the FOM of the ADC itself is *doubled*. Thus, in general, the driver and the ADC must be cooptimized.

IV. Effect of Mismatches

The effect of offset, gain and timing mismatches in an interleaved ADC can be analyzed in the time or frequency domains with sinusoidal or random inputs [1]. Each of these four approaches provides its own insight. We present a new frequency-domain perspective here [12] that reveals the mechanisms by which mismatches corrupt the analog input and exposes subtle effects that, evidently, have not been appreciated in prior work.

Let us assume that a two-channel ADC senses a random analog input, $x(t)$, with a bandwidth of $\pm f_{\mathrm{in}}$. The signal is sampled, quantized, and multiplexed, but we ignore the effect of quantization here. If the two channels are matched, their outputs can be expressed as

$$y_1(t) = x(t) \sum_k \delta(t - kT_{\mathrm{CK}}) \tag{16}$$

$$y_2(t) = x(t) \sum_k \delta\left(t - kT_{\mathrm{CK}} - \frac{T_{\mathrm{CK}}}{2}\right). \tag{17}$$

[5]In architectures that allow less than $0.5T_{\mathrm{CK}}$ for acquisition, the FOM is multiplied by a value smaller than $N/2$.

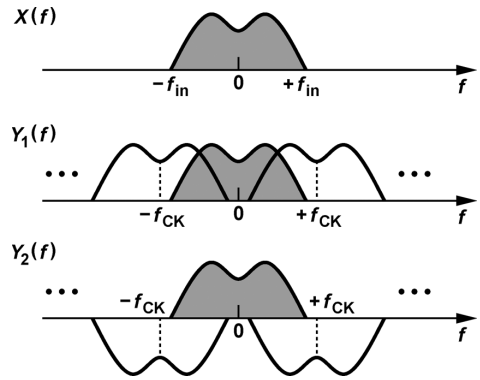


Fig. 6. Input and output spectra of a two-channel interleaved ADC.

We refer to the y_1 and y_2 samples as odd and even, respectively. In the frequency domain

$$Y_1(f) = \frac{1}{T_{\mathrm{CK}}} X(f) * \sum_k \delta(f - kf_{\mathrm{CK}}) \tag{18}$$

$$Y_2(f) = \frac{1}{T_{\mathrm{CK}}} X(f) * \sum_k \delta(f - kf_{\mathrm{CK}}) e^{-j2\pi f T_{\mathrm{CK}}}. \tag{19}$$

Shown in Fig. 6, these spectra exhibit heavy aliasing, which is eventually undone by the back-end multiplexer. Since multiplexing of discrete-time signals is equivalent to addition, the spectral copies around $\pm f_{\mathrm{CK}}$ cancel each other in $Y_1(f) + Y_2(f)$, thus yielding the original $X(f)$. Such cancellation is reminiscent of image rejection in RF receivers, pointing to the precise matching required of the channels.

A. Offset and Gain Mismatch

The effect of offset and gain mismatches has been studied extensively [1], [2], [26], [30], [27]. In the absence of the input signal, the two ADCs digitize their own DC offsets. Consequently, the multiplexed output toggles between the two offsets with a period equal to T_{CK}, thus exhibiting a tone at $f_{\mathrm{CK}} = f_S/2$. The power of this tone can be added to noise and harmonic powers [27] but, as an unmodulated sinusoid, it may not be objectionable.

The effect of gain mismatch can be readily observed from the spectra in Fig. 6. As shown in Fig. 7, the spectral copies centered at $\pm f_{\mathrm{CK}}$ in $Y_1(f)$ and $Y_2(f)$ do not completely cancel, corrupting the baseband signal by its aliased components. For a (voltage) gain mismatch of α, we can express the power of these components integrated from $-f_{\mathrm{in}}$ ($\approx -f_{\mathrm{CK}}$) to $+f_{\mathrm{in}}$($\approx +f_{\mathrm{CK}}$) as $\alpha^2 P_X$, where P_X denotes the total power of $X(f)$. The signal power in $Y_1(f) + Y_2(f)$ is equal to $4P_X$ because the desired components in $Y_1(f)$ and $Y_2(f)$ add in-phase; it follows that:

$$\mathrm{SNR} = \frac{4}{\alpha^2}. \tag{20}$$

B. Timing Mismatch

The effect of timing mismatch in interleaved samplers can be viewed as follows: if a signal, $x(t)$, is delayed by a small amount, ΔT, then $x(t + \Delta T) \approx x(t) + \Delta T dx/dt$. Thus, the

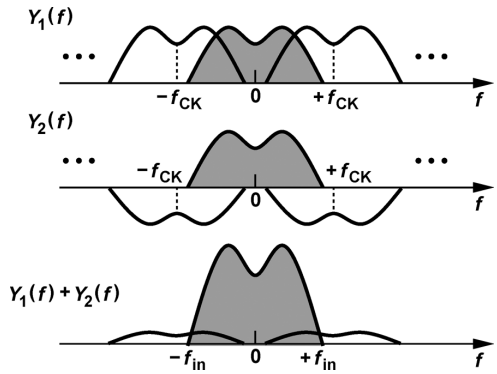


Fig. 7. Effect of gain mismatch in an interleaved ADC.

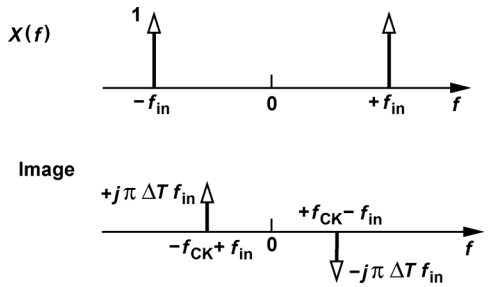


Fig. 8. Effect of timing mismatch with a sinusoidal input.

timing mismatch between two channels creates an error equal to $\Delta T dx/dt$.

Let us first consider this effect with a sinusoidal input, $x(t) = A\cos\omega_{\rm in}t$, where $\omega_{\rm in} = 2\pi f_{\rm in}$. As derived in [30], the multiplexed output can be expressed as

$$y(nT_S) = \cos\frac{\omega_{\rm in}\Delta T}{2}\cos\left(\omega_{\rm in}nT_S + \frac{\omega_{\rm in}\Delta T}{2}\right) + \sin\frac{\omega_{\rm in}\Delta T}{2}\sin\left[(\omega_{\rm in} - \pi f_S)nT_S + \frac{\omega_{\rm in}\Delta T}{2}\right] \quad (21)$$

where the second term represents an aliased component (called the "image" in [30]) that limits the SNR to

$$\text{SNR} = \frac{1}{\pi^2\Delta T^2 f_{\rm in}^2}. \quad (22)$$

Fig. 8 plots the corresponding spectrum.

The sinusoidal test entails two drawbacks. First, it overestimates the corruption by considering only the highest input frequency. A realistic broadband signal would suffer maximum error only near the edge of its bandwidth. Second, (22) does not reveal another corruption mechanism.

We now analyze the effect of timing mismatch by applying a random input signal, $x(t)$. For ease of illustration, we assume $x(t)$ has a flat spectrum from $-f_{\rm in}$ to $+f_{\rm in}$ [Fig. 9(a)]. The additive term, $\Delta T dx/dt$, translates to $\Delta T[j2\pi fX(f)]$ in the frequency domain. Thus, (18) and (19) are respectively written as

$$Y_1(f) = \frac{1}{T_{\rm CK}}X(f) * \sum_k \delta(f - kf_{\rm CK}) \quad (23)$$

$$Y_2(f) = \frac{1}{T_{\rm CK}}[X(f) + j2\pi f\Delta TX(f)] * \sum_k \delta(f - kf_{\rm CK})e^{-j\pi fT_{\rm CK}}. \quad (24)$$

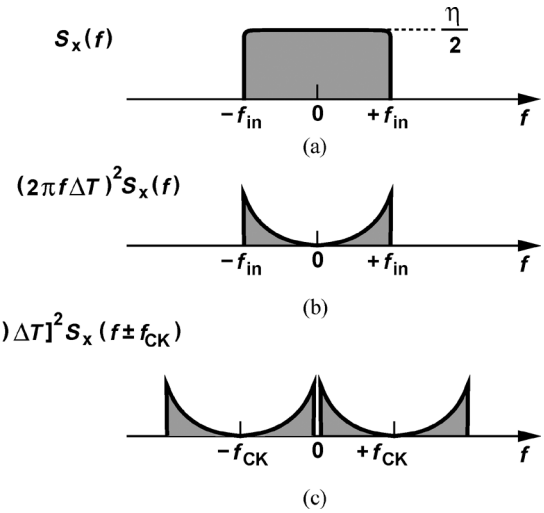


Fig. 9. Effect of timing mismatch with a random input: (a) input spectrum; (b) untranslated first-order shaped component; and (c) translated first-order shaped component.

We study $Y_2(f)$ for $k = 0$ and $k = 1$. For the former case, $\delta(f)\exp(-j\pi fT_{\rm CK}) = \delta(f)$ and hence

$$Y_{2,k=0}(f) = \frac{1}{T_{\rm CK}}[X(f) + j2\pi f\Delta TX(f)]. \quad (25)$$

We observe that $Y_{2,k=0}(f)$ contains $X(f)$ itself and a new term, $j2\pi f\Delta TX(f)$, which is the first-order shaped version of $X(f)$. Fig. 9(b) plots the spectral density of the new term, ignoring for simplicity the factor of $1/T_{\rm CK}^2$ predicted by (25).

For $k = 1$, we have $\delta(f \pm f_{\rm CK})\exp(-j\pi fT_{\rm CK}) = -\delta(f \pm f_{\rm CK})$ and hence

$$Y_{2,k=1}(f) = \frac{-1}{T_{\rm CK}}[X(f \pm f_{\rm CK}) + j2\pi(f \pm f_{\rm CK})\Delta TX(f \pm f_{\rm CK})]. \quad (26)$$

The main term, $-X(f \pm f_{\rm CK})$, cancels a similar term in $Y_1(f)$, while the shaped component aliases into the signal band [Fig. 9(c)].

It is instructive to reexamine the results obtained with a sinusoidal input in light of the mechanisms revealed by the random input. The component at $f_{\rm CK} - f_{\rm in}$ in Fig. 8 corresponds to the aliased term $2\pi(f \pm f_{\rm CK})\Delta TX(f \pm f_{\rm CK})$ in $Y_{2,k=1}(f)$. But what in Fig. 8 corresponds to the unaliased term $2\pi f\Delta TX(f)$ in $Y_{2,k=0}(f)$? Let us expand (21)

$$y(nT_S) = \frac{1}{2}\cos(\omega_{\rm in}nT_S + \omega_{\rm in}\Delta T) + \frac{1}{2}\cos(\omega_{\rm in}nT_S) + \sin\frac{\omega_{\rm in}\Delta T}{2}\sin\left[(\omega_{\rm in} - \pi f_S)nT_S + \frac{\omega_{\rm in}\Delta T}{2}\right]. \quad (27)$$

The first term on the right-hand side is simply a time-shifted copy of the second, representing the effect of sampling with a timing mismatch of ΔT. Assuming $\omega_{\rm in}\Delta T \ll 1$ rad, we expand the first term, obtaining

$$y(nT_S) \approx \cos(\omega_{\rm in}nT_S) - \frac{\omega_{\rm in}\Delta T}{2}\sin(\omega_{\rm in}nT_S) + \frac{\omega_{\rm in}\Delta T}{2}\sin\left[(\omega_{\rm in} - \pi f_S)nT_S + \frac{\omega_{\rm in}\Delta T}{2}\right]. \quad (28)$$

The second term arises from the derivative of the signal, but unlike the third term, it has not experienced frequency translation. Hiding behind the main component, this term has evidently not been considered in prior work because it is indistinguishable from the main signal for a sinusoidal input. The equal amplitudes of the second and third terms in (28) suggest that neglecting the second term underestimates the corruption by 3 dB.

That the derivative of a random signal corrupts the signal itself can be intuitively explained as follows. The correlation between $x(t)$ and dx/dt can be expressed as

$$R_{x,dx/dt} = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{+T/2} x(t) \frac{dx}{dt} dt \tag{29}$$

$$= \lim_{T\to\infty} \frac{1}{T} \left[\frac{1}{2} x^2(+T/2) - \frac{1}{2} x^2(-T/2) \right]. \tag{30}$$

Interestingly, so long as $x(t)$ is bounded, $R_{x,dx/dt}$ is zero, indicating that dx/dt has no resemblance to $x(t)$ and hence completely corrupts the signal if added to it.

We now wish to formulate the SNR for a finite ΔT and a random input signal having a flat spectrum from $-f_{\rm in}$ to $+f_{\rm in}$. Let us first find the SNR due to only the aliased spectrum of Fig. 9(c) and compare the result with (22). Integrating this spectrum from $-f_{\rm in}$ to $+f_{\rm in}$, we have

$$P_{n1} = \int_{-f_{\rm in}}^{+f_{\rm in}} (2\pi f \Delta T)^2 S_X(f) df \tag{31}$$

$$= \frac{4}{3} \pi^2 \eta \Delta T^2 f_{\rm in}^3. \tag{32}$$

The desired signal power in Fig. 9(a) after multiplexing is equal to

$$P_{\rm sig} = 4 \int_{-f_{\rm in}}^{+f_{\rm in}} \frac{\eta}{2} df \tag{33}$$

where the factor of 4 accounts for the two copies of $X(f)$ in $Y_1(f)$ and $Y_2(f)$ (Fig. 6). It follows that:

$$\frac{P_{\rm sig}}{P_{n1}} = \frac{3}{\pi^2 \Delta T^2 f_{\rm in}^2}. \tag{34}$$

As expected, the SNR is higher in this case because components lower than $f_{\rm in}$ create less noise. The factor of 3 is reminiscent of other situations where a "distributed" quantity is approximated by a lumped quantity, e.g., the noise power due to the distributed gate resistance of a MOSFET [31] or the input impedance of an RC ladder [32]. In essence, (22) assumes that the signal power distributed from 0 to $f_{\rm in}$ is lumped into an impulse at $f_{\rm in}$. (The factor of 3 disparity has also been recognized in [28]).

In the next step, we also include the power of the unaliased spectrum in Fig. 9(b)

$$\mathrm{SNR} = \frac{3}{2\pi^2 \Delta T^2 f_{\rm in}^2}. \tag{35}$$

This result is still 1.76 dB higher than that predicted by (22).

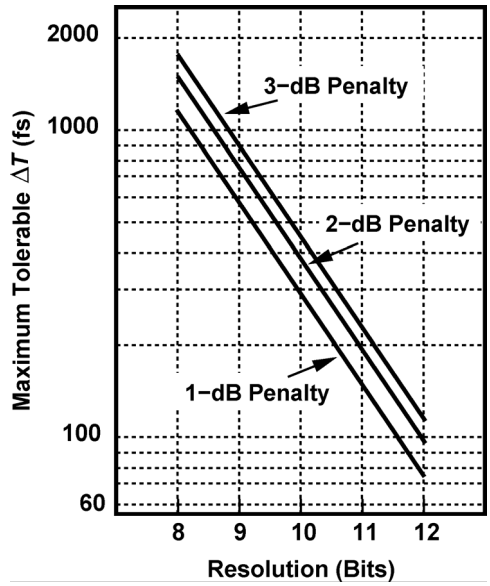


Fig. 10. Maximum tolerable timing mismatch for different SNR penalties for $f_{\rm in} = 500$ MHz.

The performance degradation due to the timing mismatch must be considered in conjunction with the quantization noise (and, eventually, kT/C noise). For an input sinusoid, we have

$$\mathrm{SNR}_{\rm tot} = \frac{V_0^2/2}{\frac{1}{12}\left(\frac{2V_0}{2^M}\right)^2 + \pi^2 \Delta T^2 f_{\rm in}^2}. \tag{36}$$

Fig. 10 plots the maximum tolerable value of ΔT as a function of M for different SNR penalties. An input frequency of 500 MHz is assumed to underscore the extremely tight phase matching necessary in today's designs.

V. Proposed Timing Mismatch Calibration

The interleaving issues described in the previous sections call for additional circuit and architecture techniques. For example, the high input capacitance and the timing mismatch can be avoided through the use of a full-speed front-end sampler [17]. In this case, however, this sampler must be followed by a fast-settling buffer to drive the channels. For the other issues, namely, the interchannel mismatches, numerous methods have been proposed [2]–[11]. While offset and gain mismatches can be readily measured and corrected in the digital domain [2], the problem of timing mismatch presents its own challenges. Among the copious approaches addressing this effect [2]–[11], very few have been experimentally verified for resolutions around 10 bits and high sampling rates; for example, the ADC in [11] operates at $f_S = 2.8$ GHz but it requires an extra channel for timing mismatch detection. We propose herein a new technique [12] and demonstrate its efficacy by experimental results.

The task of timing mismatch calibration consists of two functions, namely, detection and correction, each of which can be performed in the analog or digital domain. Detection in the analog domain is prone to the mismatches of the measurement

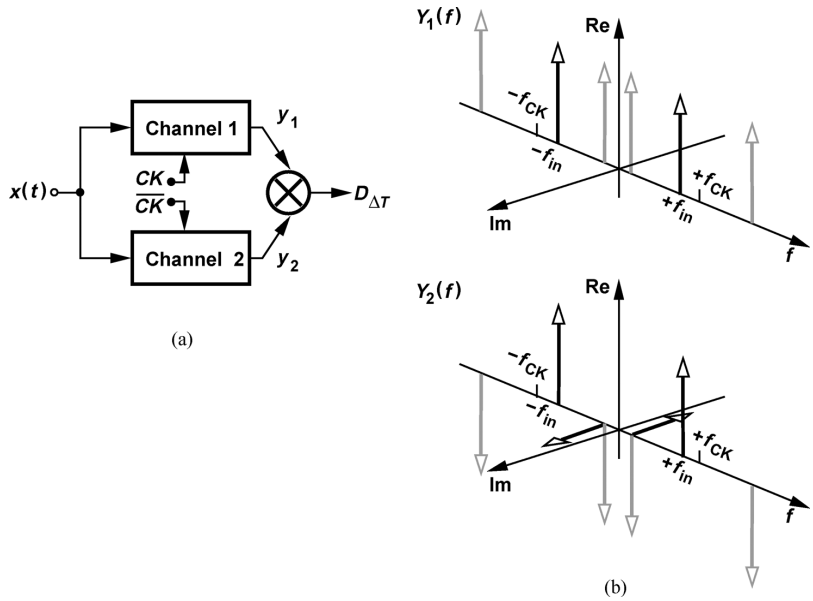


Fig. 11. (a) Attempt to compute timing error by mixing, and (b) corresponding spectra for $x(t) = A\cos(2\pi f_{\rm in}t)$.

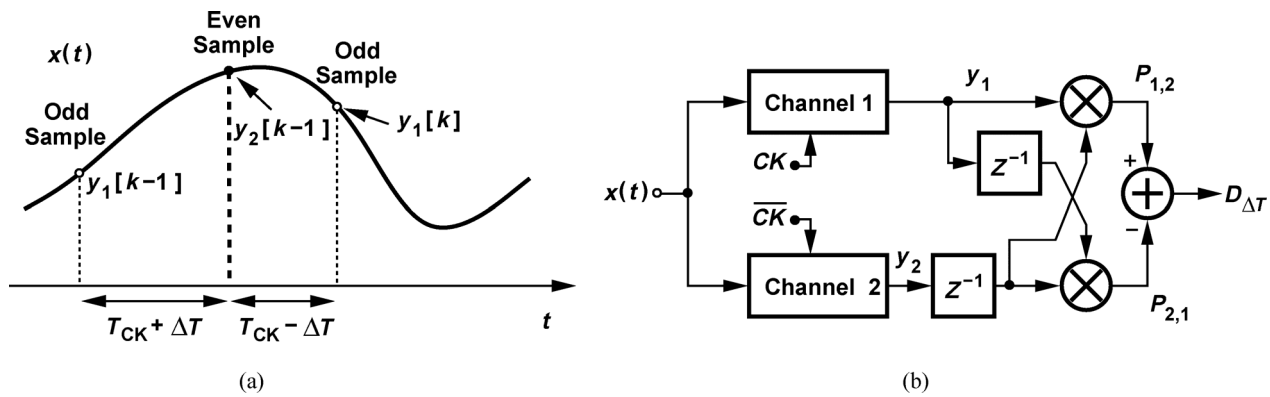


Fig. 12. (a) Waveform illustrating the effect of timing mismatch on $y_1[k-1]y_2[k-1]$ and $y_2[k-1]y_1[k]$, and (b) timing mismatch detection topology.

circuitry itself (e.g., a mixer) and hence impractical for high resolutions. Correction of mismatches can be realized in the analog domain [3], [4], [10], [11] or in the digital domain [2].

A. Proposed Detection Method

In analogy with phase detection in PLLs and RF circuits, we may consider digital mixing as a means of timing mismatch detection. Illustrated in Fig. 11(a), the idea is to multiply the outputs of two interleaved channels and utilize the product's average, $D_{\Delta T}$, as a measure of the mismatch. Unfortunately, this approach fails at least in some cases. For a sinusoidal input, $x(t) = A\cos(2\pi f_{\rm in}t)$, the timing mismatch produces an error given by $\Delta T dx/dt = -\Delta T A(2\pi f_{\rm in})\sin(2\pi f_{\rm in}t)$, which is then translated by $f_{\rm CK}$. The two channels' output spectra thus emerge as in Fig. 11(b), where the image is shown along the imaginary axis and the gray impulses represent the translated spectra as in Fig. 6. We observe that the mixing of these outputs cannot generate a dc term proportional to the image amplitude. This is because the image components carrying the mismatch information [the black imaginary impulses in $Y_2(f)$] are orthogonal to the gray impulses in $Y_1(f)$, yielding a zero average for their mixing result.

Let us now consider forming *two* products: the product of an odd sample [from $y_1(t)$] and the next even sample [from $y_2(t)$], and the product of an even sample and the next odd sample. As illustrated in Fig. 12(a), the time difference between an odd sample and the next even sample is, e.g., greater than $T_{\rm CK}$, thereby skewing the long-term average of their product. Similarly, the time difference between an even sample and the next odd sample is, e.g., less than $T_{\rm CK}$, yielding an *opposite* skew in their product's average. The difference between the two averages is thus proportional to ΔT.

The foregoing intuition can also be confirmed mathematically. Let us choose the time origin such that three consecutive samples of the input signal can be expressed as $x[-(T_{\rm CK}+\Delta T)]$, $x(0)$, and $x(T_{\rm CK}-\Delta T)$. We recognize that the time average of $x[-(T_{\rm CK}+\Delta T)]x(0)$ is in fact the autocorrelation of the input, $R_x(\tau)$, evaluated at $-(T_{\rm CK}+\Delta T)$, i.e., $R_x[-(T_{\rm CK}+$

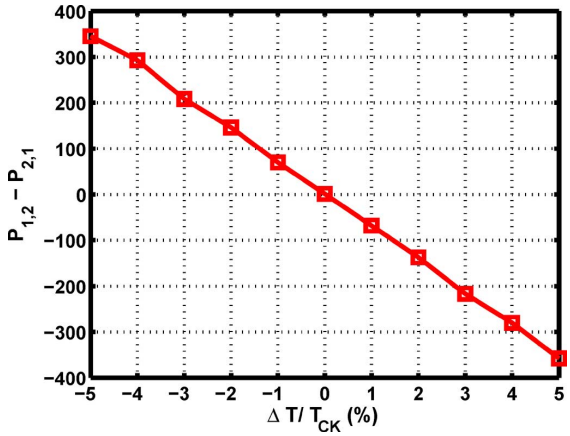


Fig. 13. Simulated average difference between the two products as a function of timing error (arbitrary vertical unit).

$\Delta T)]$. Similarly, the time average of $x(0)x(T_{\rm CK}-\Delta T)$ is equal to $R_x(T_{\rm CK}-\Delta T)$. Thus, the average value of $D_{\Delta T}$ is given by

$$\overline{D_{\Delta T}} = R_x(T_{\rm CK}-\Delta T) - R_x[-(T_{\rm CK}+\Delta T)]. \quad (37)$$

Since the autocorrelation is an even function, $R_x[-(T_{\rm CK}+\Delta T)] = R_x(T_{\rm CK}+\Delta T)$ and hence

$$\overline{D_{\Delta T}} = R_x(T_{\rm CK}-\Delta T) - R_x(T_{\rm CK}+\Delta T). \quad (38)$$

For a small ΔT, this difference reduces to

$$\overline{D_{\Delta T}} \approx -2\Delta T\frac{{\rm dR}_x}{d\tau}, \quad (39)$$

where $dR_x/d\tau$ is computed at $\tau = T_{\rm CK}$.

In order to implement the above detection method, we note from Fig. 12(a) that two products of the form $y_2[k-1]y_1[k-1]$ and $y_1[k]y_2[k-1]$ are necessary. Depicted in Fig. 12(b) is the digital implementation, where $P_{1,2} = y_1[k]y_2[k-1]$ and $P_{2,1} = y_2[k-1]y_1[k-1]$.

The proposed timing mismatch detection technique has been studied with sinusoidal and random inputs. Fig. 13 plots the simulated average of $P_{1,2}-P_{2,1} = D_{\Delta T}$ as a function of $\Delta T/T_{\rm CK}$ for a random input whose bandwidth is limited to $\pm f_{\rm CK}$. We observe that the difference varies monotonically and changes sign as ΔT crosses zero, serving as a measure of the mismatch.

The digital implementation shown in Fig. 12(b) incorporates two registers (for the z^{-1} operation), two multipliers, and one adder, potentially consuming a high power. Fortunately, $D_{\Delta T}$ need not be updated at the full clock rate. In other words, the delay, multiply, and add operations can be repeated much less frequently than the clock speed so as to minimize the power drawn by the logic. The update rate must nonetheless be fast enough to reflect drifts in ΔT due to temperature variations.

Fig. 14(a) illustrates a possible realization of this approach. Here, the ΔT computation of Fig. 12(b) is followed by a register, a DAC, and an analog variable-delay line (VDL), which adjusts the delay of the channel 2 clock. In the behavioral simulation of this architecture, a timing mismatch of 3% is included between the two channels. The ΔT computation proceeds for 8000 clock cycles at full speed, updating $D_{\Delta T}$, and then rests for 10^6 cycles. The register, the DAC, and the VDL therefore freeze the updated delay until the next computation cycle. This compute-and-rest sequence is repeated periodically, producing the behavior shown in Fig. 14(b) for the timing mismatch as a function of time. (The time difference between the iterations is 10^6 cycles.) The loop drives the mismatch toward small values, while the power consumption of the logic is reduced by a factor of $10^6/8000 = 125$.

The choice of compute and rest cycles is somewhat flexible. For example, a clock rate around 1 GHz allows a rest period of $10^6 \times 1$ ns $= 1$ ms, accommodating typical temperature changes. For lower clock speeds or faster temperature variations, the number of cycles for the rest period can be reduced. Note that the calibration operates in the background while assuming that the offset and gain mismatches have already been removed by other means.

B. Proposed Correction Method

The timing mismatch can be corrected in the digital domain by means of a finite-impulse-response (FIR) filter [2] or in the analog domain by means of a variable delay line [3], [4], [10], [11]. The plots of Fig. 10 demand a correction step of less than 300 fs if a 500-MHz sinusoid must be digitized with no more than 1-dB SNR penalty. We must therefore decide between a 500-MHz FIR filter that is long enough to yield such a small step and an analog delay line whose phase noise is low enough not to degrade the SNR. This work employs the latter and targets a step of 40 fs. The challenge lies in the design of one or more stages whose delay can be adjusted in such minute steps.

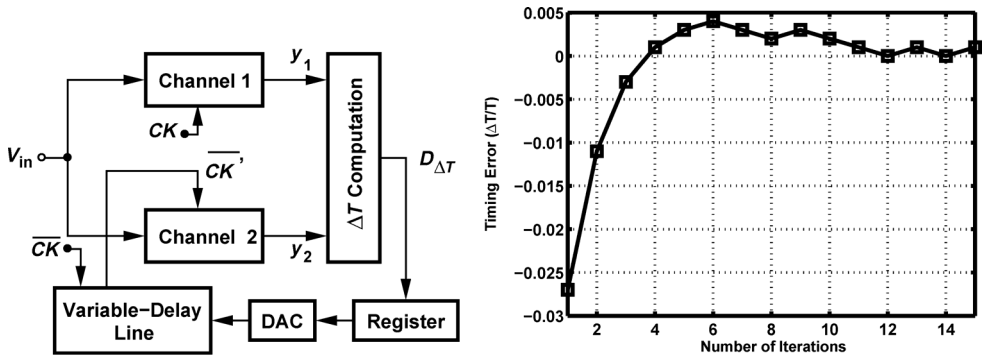
In addition to the minimum correction step, the total timing mismatch must also be estimated so as to determine the overall resolution of the variable delay line. Monte Carlo simulations of the clock paths in this design suggest a maximum mismatch of 2 ps, calling for 2 ps/40 fs $= 50$ correction steps and hence, a resolution of 6 bits. (The signal driving the VDL is retimed by a master clock so as to avoid accumulation of mismatches due to the nonoverlap clock generators and the buffers.) The sign of the mismatch can be accommodated if each channel incorporates a VDL, allowing differential control of the delay.

Fig. 15 shows the implementation of the variable-delay line. Here, transistor M_2 is degenerated by a fixed on-resistance, that of M_3, in parallel with a variable on-resistance, that of M_4. If M_3 is chosen substantially stronger than M_4, the delay can be adjusted in fine steps. The gate voltage of M_4 is driven by a 6-bit resistor ladder.

VI. Experimental Results

The detection and correction techniques proposed in Section V have been experimentally verified on a two-channel interleaved ADC. Fig. 16 shows the die photograph.[6] Each channel employs the 10-bit ADC design reported in [33] along with an on-chip variable delay line, which consists of the 6-bit stage shown in Fig. 15 preceded by another stage with a 1-bit coarse control. The outputs of the two channels are multiplexed, downsampled, and sent off-chip to Matlab, where the detection method of Fig. 12(b) is applied to compute the timing

[6]This die includes four channels but only two have been used for this demonstration.



(a) (b)

Fig. 14. (a) Proposed background calibration method, and (b) its convergence behavior.

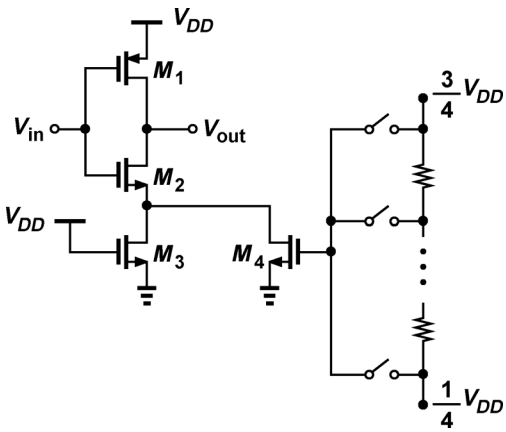


Fig. 15. Proposed variable delay line.

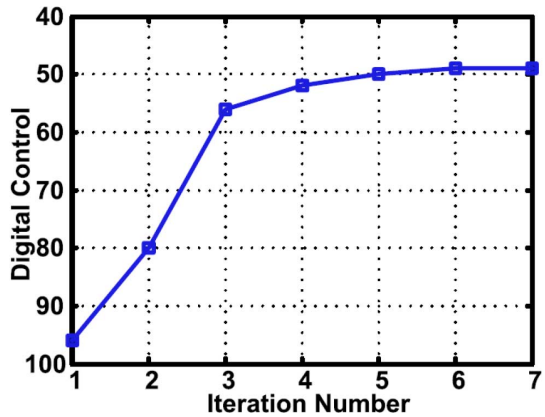


Fig. 17. Measured VDL input code during convergence.

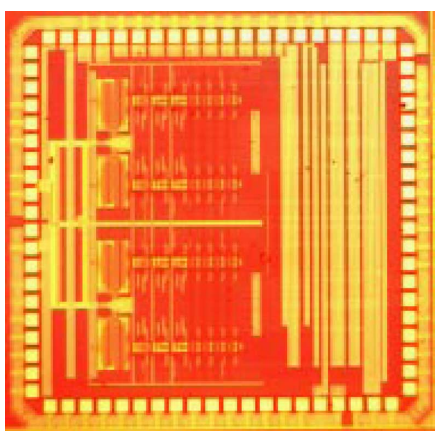


Fig. 16. Prototype ADC die photograph.

mismatch.[7] The result then returns to the chip through a serial bus and adjusts the variable-delay lines. The measured results are obtained using the compute-and-rest method depicted in Fig. 14(a) with a sinusoidal input. The prototype operates with a 1.2-V supply, consuming the same power per channel as that in [33] plus 1 mW for each VDL. The simulated rms jitter of the VDL is about 25 fs.

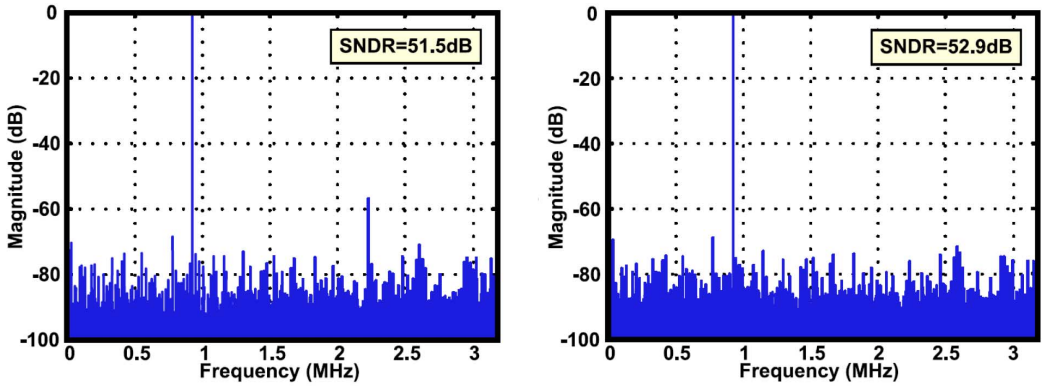
[7]Offset and gain mismatches are determined and corrected manually in Matlab.

Fig. 17 plots the VDL digital input code as the loop converges. The iteration number refers to each compute-and-rest cycle (about 13 μs). In this test, $f_{\rm in} = 396$ MHz and $f_S = 800$ MHz. Fig. 18 shows the output spectra before and after calibration, demonstrating that the image can be reduced from -56 dB to about -75 dB. Used for only confirming the proposed calibration technique, the overall ADC is not optimized for SNDR or speed and hence has a higher FOM than the single-channel prototype in [33].

Fig. 19 repeats the above experiment for $f_{\rm in} = 596$ MHz and $f_S = 800$ MHz so as to verify the calibration above the Nyquist rate. The image falls from -52 dB to about -75 dB.

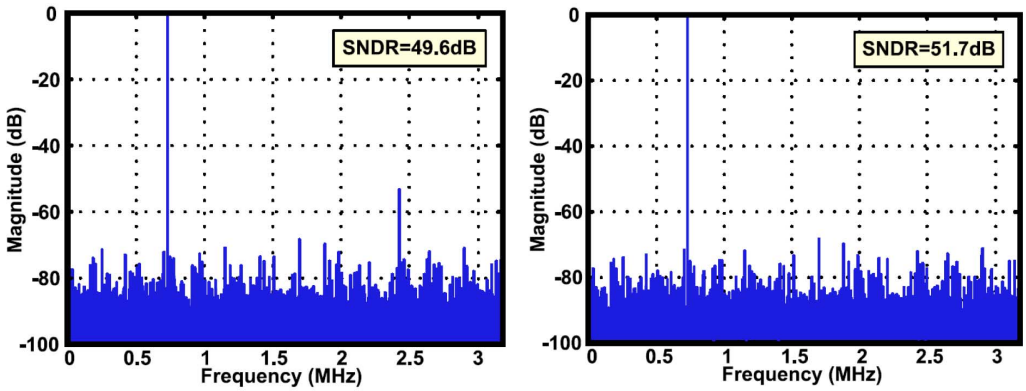
VII. Conclusion

In addition to raising the conversion speed, interleaved ADCs can also lower the power consumption and the metastability error rates. This paper identifies weakly-scalable timing overheads in ADCs that tend to degrade the FOM at higher speeds but become less significant if interleaving is employed. Moreover, an upper bound on the performance of A/D interfaces due to kT/C noise is derived that reveals the power penalty due to driver circuits preceding ADCs, and more seriously, preceding interleaved systems. A frequency-domain analysis of interchannel mismatches provides additional insight into the signal corruption mechanisms. A background timing mismatch



(a) (b)

Fig. 18. Measured output spectra for $f_{\rm in} = 396$ MHz and $f_S = 800$ MHz (a) before, and (b) after calibration. (Due to downsampling by a factor of 125, the Nyquist band maps to 0 to 3.2 MHz, $f_{\rm in}$ to 0.95 MHz, and image to 2.25 MHz.).



(a) (b)

Fig. 19. Measured output spectra for $f_{\rm in} = 596$ MHz and $f_S = 800$ MHz (a) before, and (b) after calibration. (Due to downsampling by a factor of 125, the Nyquist band maps to 0 to 3.2 MHz, $f_{\rm in}$ to 0.75 MHz, and the image to 2.45 MHz.)

calibration method is also introduced that reduces the image to -75 dB for input frequencies exceeding 500 MHz.

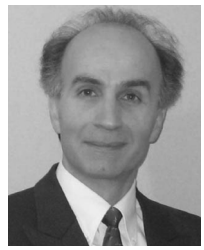
ACKNOWLEDGMENT

The author would like to thank H. Wei and W. Chiang for their contributions to the proposed calibration method. He also gratefully acknowledges TSMC's University Shuttle Program for chip fabrication. This work was supported by the DARPA HEALICS Program, Realtek Semiconductor, and Pullman Lane Productions.

REFERENCES

[1] W. C. Black and D. A. Hodges, "Time-interleaved converter arrays," *IEEE J. Solid-State Circuits*, vol. 15, no. 12, pp. 1022–1029, Dec. 1980.
[2] S. M. Jamal *et al.*, "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.
[3] C.-C. Huang, C.-Y. Wang, and J.-T. Wu, "A CMOS 6-bit 16-GS/s time-interleaved ADC using digital background calibration techniques," *IEEE J. Solid-State Circuits*, vol. 46, no. 4, pp. 848–858, Apr. 2011.
[4] M. El-Chammas and B. Murmann, "A 12-GS/s 81-mW 5-bit time-interleaved flash ADC with background timing skew calibration," *IEEE J. Solid-State Circuits*, vol. 46, no. 4, pp. 838–847, Apr. 2011.
[5] J. A. McNeill *et al.*, "Split ADC calibration for all-digital correction of time-interleaved ADC errors," *IEEE Tran. Circuits Syst. II*, vol. 56, no. 5, pp. 344–348, May 2009.
[6] A. Haftbaradaran and K. W. Martin, "A sample-time error compensation technique for time-interleaved ADC systems," in *Proc. CICC*, Sept. 2007, pp. 341–344.
[7] J. Elbornsson, F. Gustafsson, and J.-E. Eklund, "Blind equalization of time errors in a time-interleaved ADC system," *IEEE Tran. Signal Process.*, vol. 53, no. 4, pp. 1413–1424, Apr. 2005.
[8] S. Huang and C. Levy, "Adaptive blind calibration of timing offset and gain mismatch for two-channel time-interleaved ADCs," *IEEE Tran. Circuits Syst. I*, vol. 53, no. 6, pp. 1278–1288, Jun. 2006.
[9] V. Divi and G. W. Wornell, "Blind calibration of timing skew in time-interleaved analog-to-digital converters," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 6, pp. 509–522, Jun. 2009.
[10] C.-Y. Wang and J.-T. Wu, "A background timing-skew calibration technique for time-interleaved analog-to-digital converters," *IEEE Trasn. Circuits Syst. II*, vol. 53, no. 4, pp. 299–303, Apr. 2006.
[11] D. Stepanovic and B. Nikolic, "A 2.8-GS/s 44.6-mW time-interleaved ADC achieving 50.9 SNDR and 3-dB effective resolution bandwidth of 1.5 GHz in 65-nm CMOS," in *Proc. Symp. VLSI Circuits Dig. Tech. Papers*, Jun. 2012, pp. 84–85.
[12] B. Razavi, "Problem of timing mismatch in interleaved ADCs," presented at the IEEE CICC, Sep. 2012.
[13] L. Sumanen, M. Waltari, and K. Halonen, "A 10-bit 200-MS/s CMOS parallel pipeline A/D converter," *IEEE J. Solid-State Circuits*, vol. 46, no. 7, pp. 1048–1055, Jul. 2001, vol. 36.
[14] B. Verbruggen, M. Iriguchi, and J. Craninckx, "A 1.7-mW 11-b 250-MS/s 2x interleaved fully dynamic pipelined SAR ADC in 40-nm digital CMOS," in *Proc. ISSCC Dig. Tech. Papers*, Feb. 2012, pp. 466–467.
[15] J. Montanaro *et al.*, "A 160-MHz, 32-b, 0.5-W CMOS RISC microprocessor," *IEEE J. Solid-State Circuits*, vol. 31, no. 11, pp. 1703–1714, Nov. 1996.
[16] P. Lim and B. A. Wooley, "An 8-bit 200-MHz BiCMOS comparator," *IEEE J. Solid-State Circuits*, vol. 25, no. 2, pp. 192–199, Feb. 1990.

[17] C. C. Hsu *et al.*, "An 11-b 800-MS/s time-interleaved ADC with digital background calibration," in *Proc. ISSCC Dig. Tech. Papers*, Feb. 2007, pp. 464–465, (also Slide Supplement).
[18] Z. Cao, S. Yan, and Y. Li, "A 32-mW 1.25-GS/s 6-b 2-b/step SAR ADC in 0.13 μm CMOS," in *Proc. ISSCC Dig. Tech. Papers*, Feb. 2008, pp. 542–543, (also Visual Supplement).
[19] iB. Zojer *et al.*, "A 6-Bit/200-MHz full Nyquist A/D converter," *IEEE J. Solid-State Circuits*, vol. 20, no. 6, pp. 780–786, Jun. 1985.
[20] S. Hashemi and B. Razavi, "A 10-Bit 1-GS/s CMOS ADC with FOM=70 fJ/Conversion," presented at the IEEE CICC, Sep. 2012.
[21] J.-S. Choi and K. Lee, "Design of CMOS tapered buffer for minimum power-delay product," *IEEE J. Solid-State Circuits*, vol. 29, no. 9, pp. 142–1145, Sep. 1994.
[22] H. Wei *et al.*, "A 0.024-mm^2 8-b 400-MS/s SAR ADC with 2-b/cycle and resistive DAC in 65-nm CMOS," in *Proc. ISSCC Dig. Tech. Papers*, Feb. 2011, pp. 188–189.
[23] J.-M. Chou *et al.*, "Phase averaging and interpolation using resistor strings or resistor rings for multi-phase clock generation," *IEEE Trans. Circuits Syst, I*, vol. 53, no. 5, pp. 984–991, May 2006.
[24] J. Guerber *et al.*, "A 10-b ternary SAR ADC with quantization time information utilization," *IEEE J. Solid-State Circuits*, vol. 47, no. 11, pp. 2604–2613, Nov. 2012.
[25] C.-Y. Chen *et al.*, "A 12-bit 3-GS/s pipeline ADC with 0.4 mm^2 and 500 mW in 40-nm digital CMOS," *IEEE J. Solid-State Circuits*, vol. 4, no. 4, pp. 1013–1021, Apr. 2012.
[26] Y. C. Jenq, "Digital spectra of nonuniformly sampled signals: Fundamentals and high-speed waveform digitizers," *IEEE Trans. Instrum. Meas.*, vol. 37, no. 6, pp. 245–251, Jun. 1988.
[27] N. Kurosawa *et al.*, "Explicit analysis of channel mismatch effects in time-interleaved ADC systems," *IEEE Tran. Circuits Syst. I*, vol. 38, no. 3, pp. 261–271, Mar. 2001.
[28] M. El-Chammas and B. Murmann, "General analysis on the impact of phase-skew in time-interleaved ADCs," *IEEE Tran. Circuits Syst. I*, vol. 56, no. 5, pp. 902–910, May 2009.
[29] Y. T. Wang and B. Razavi, "An 8-bit 150-MHz CMOS A/D converter," *IEEE J. Solid-State Circuits*, vol. 35, no. 3, pp. 308–317, Mar. 2000.
[30] S. M. Jamal *et al.*, "Calibration of sample-time error in a two-channel time-interleaved analog-to-digital converter," *IEEE Trans. Circuits Syst. ,I*, vol. 51, no. 1, pp. 130–139, Jan. 2004.
[31] B. Razavi, R. H. Yan, and K. F. Lee, "Impact of distributed gate resistance on the performance of MOS devices," *IEEE Trans. Circuits Syst., I*, vol. 41, no. 11, pp. 750–754, Nov. 1994.
[32] B. Razavi, *RF Microelectron.*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2011.
[33] B. Sahoo and B. Razavi, "A 10-Bit 1-GHz 33-mW CMOS ADC," in *Proc. Symp. VLSI Circuits Dig. Of Tech.*, Jun. 2012, pp. 30–31.



Behzad Razavi (F'03) received the B.S. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 1985, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, USA, in 1988 and 1992, respectively.

He was with AT&T Bell Laboratories and Hewlett-Packard Laboratories until 1996. Since 1996, he has been Associate Professor and subsequently Professor of electrical engineering at University of California, Los Angeles, CA, USA. His current research includes wireless transceivers, frequency synthesizers, phase-locking and clock recovery for high-speed data communications, and data converters. He was an Adjunct Professor at Princeton University from 1992 to 1994, and at Stanford University in 1995. He served on the Technical Program Committees of the International Solid-State Circuits Conference (ISSCC) from 1993 to 2002 and VLSI Circuits Symposium from 1998 to 2002. He has served as an IEEE Distinguished Lecturer, and is the author of "Principles of Data Conversion System Design" (IEEE Press, 1995), "RF Microelectronics" (Prentice Hall, 1998, 2012) (translated to Chinese, Japanese, and Korean), "Design of Analog CMOS Integrated Circuits" (McGraw-Hill, 2001) (translated to Chinese, Japanese, and Korean), "Design of Integrated Circuits for Optical Communications" (McGraw-Hill, 2003, Wiley, 2012), and "Fundamentals of Microelectronics" (Wiley, 2006) (translated to Korean and Portuguese).

Dr. Razavi received the Beatrice Winner Award for Editorial Excellence at the 1994 ISSCC, the Best Paper Award at the 1994 European Solid-State Circuits Conference, the best panel award at the 1995 and 1997 ISSCC, the TRW Innovative Teaching Award in 1997, the best paper award at the IEEE Custom Integrated Circuits Conference in 1998, and the McGraw-Hill First Edition of the Year Award in 2001. He was the corecipient of both the Jack Kilby Outstanding Student Paper Award and the Beatrice Winner Award for Editorial Excellence at the 2001 ISSCC. He received the Lockheed Martin Excellence in Teaching Award in 2006, the UCLA Faculty Senate Teaching Award in 2007, and the CICC Best Invited Paper Award in 2009 and 2012. He was the corecipient of the 2012 VLSI Circuits Symposium Best Student Paper Award. He was also recognized as one of the top 10 authors in the 50-year history of ISSCC. He received the IEEE Donald Pederson Award in Solid-State Circuits in 2012. He has also served as Guest Editor and Associate Editor of the IEEE JOURNAL OF SOLID-STATE CIRCUITS, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS, and the *International Journal of High Speed Electronics*. He is also the Editor of *Monolithic Phase-Locked Loops and Clock Recovery Circuits* (IEEE Press, 1996), and *Phase-Locking in High-Performance Systems* (IEEE Press, 2003).