BACKGROUND CALIBRATION OF TIMING SKEW IN TIME-INTERLEAVED A/D CONVERTERS

Boris Murmann[1] and *Manar El-Chammas*[2]

[1]Department of Electrical Engineering, Stanford University, Stanford, CA, USA
[2]Texas Instruments, Dallas, Texas, USA
Emails: murmann@stanford.edu, manar@ti.com

ABSTRACT

This paper describes theoretical details of a calibration method that addresses the problem of timing skew in time-interleaved A/D converters. The algorithm utilizes an extra ADC outside the main signal path to estimate cross correlations between each sub-channel and samples taken along the calibration ADC's ideal timing grid. Using the cross correlations, an iterative algorithm adjusts the sampling time instants in the main signal path to maximize the converter's signal-to-noise ratio. The proposed scheme was implemented in hardware and improved the signal-to-noise ratio of an experimental 5-bit, 12-GS/s ADC by about 12 dB for inputs near Nyquist.

Keywords— A/D conversion, time-interleaving, background calibration, cross correlation

1. INTRODUCTION

Time-interleaving [1] has evolved as a popular technique for increasing the conversion rate of analog-to-digital converters (ADCs). However, time-interleaved ADCs suffer from errors related to mismatch between the individual channels. In practice, the primary issue in high-speed designs such as [2], [3] is timing skew, which results in non-uniform input sampling. The impact of timing skew increases with input frequency, and therefore tends to overshadow other nonidealities for broadband inputs.

The scheme discussed in this paper was designed to overcome the effects of timing skew through the use of a statistics-based background calibration algorithm. As described in [2], this algorithm can be used in a practical circuit realization to adjust digitally controlled delay lines toward minimum timing skew between the interleaved channels. The purpose of this paper is to provide theoretical details of the algorithm that were not covered in [2].

2. OVERVIEW AND PROBLEM DEFINITION

The general architecture of the ADC considered in this work (including the skew estimation circuitry), is shown in Fig. 1.

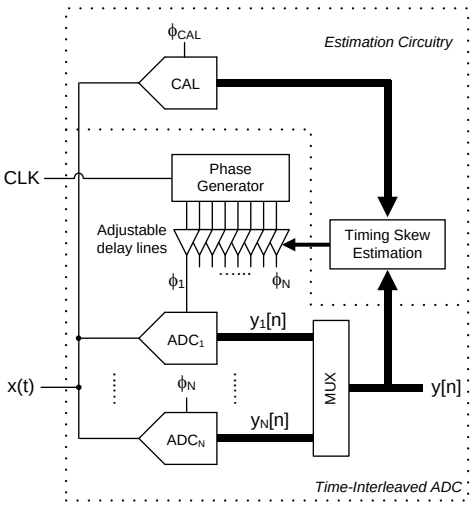

Fig. 1. ADC architecture.

In this structure, the aggregate throughput is N times the sample rate of the individual sub-ADCs ($ADC_1 \ldots ADC_N$). The sampling period of the overall interleaved ADC and the N sub-ADCs are T_s and NT_s, respectively. The sampling edges of two consecutive clocks (ϕ_i and ϕ_{i+1}) are ideally offset by T_s such that the input signal is uniformly sampled. Thus, the output of the i^{th} sub-ADC is

$$y_i[n] = x(t - \{nN + i\}T_s) \qquad (1)$$

Multiplexing the sub-ADC outputs results in the ideal output of the overall converter array

$$y[n] = x(nT_s) \qquad (2)$$

With mismatches in the signal and clock paths, (1) becomes

$$y_i[n] = G_i x(t - \{nN + i\}T_s - \tau_i) + o_i \qquad (3)$$

where G_i, o_i and τ_i are the sub-ADCs' gain, offset, and timing skew, respectively. Due to mismatch in these parameters, the time-interleaved ADC output no longer reduces to (2) in practice.

Of the three errors considered here, the impact of timing skew is the only one that increases with input frequency, and is thus often the main issue in high speed designs. This

can be seen qualitatively from Fig. 2. Higher frequency input signals have a larger derivative (d*x*/d*t*) and therefore cause a larger sampling error. Quantitative relationships between timing skew and ADC performance can be derived for wide-sense stationary and wide-sense cyclostationary signals [4] and relate the autocorrelation of the input signal to the ADC's signal-to-noise ratio (SNR). For the design of [2], it was found that the maximum tolerable skew is less than 1 ps, which is extremely difficult to achieve without a suitable calibration scheme.

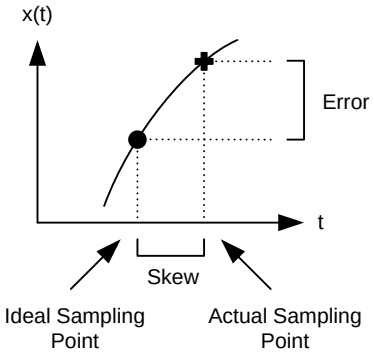

Fig. 2. Sampling error as a result of timing skew.

There are two main options for compensating the effects of timing skew. The first operates in the digital domain, employing a filter at the outputs of the sub-ADCs [5], [6]. Unfortunately, this approach typically mandates a large power overhead due to the complexity of the required filters. The second option follows a mixed-signal approach (see Fig. 1), which was preferred in the implementation of [2]. Here, a digital skew estimation block is used to adjust delay lines toward the optimal sampling clock timing in each channel. A potential disadvantage of this approach is that the employed delay lines introduce additional random jitter in the clock paths, which must be considered in the design phase.

In principle, the required parameters for skew adjustment could be measured at system start-up using an input training sequence. However, in presence of temperature variations, such an approach is rarely practical. Instead, it is preferable to perform a blind background calibration of the relevant parameters without interrupting normal ADC operation [5], [7]. In the architecture of Fig. 1, this is made possible with the addition of a calibration ADC channel labeled "CAL". As illustrated in Fig. 3, this extra channel is clocked such that its sampling instants round-robin through the nominal sampling times of each sub-ADC. This is achieved by choosing a calibration frequency of f_s/K, and K is such that the greatest common divisor between N and K is one. This timing allows us to compute the cross correlations between the outputs of the calibration ADC and each sub-ADC channel. This information can in turn be used to iteratively adjust the timing skews toward their ideal value, as discussed in the next section.

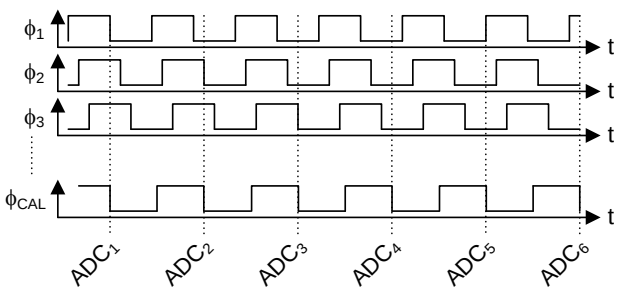

Fig. 3. Sample timing diagram for the calibration clock and sub-ADC clocks, assuming an interleaving factor of $N = 8$.

3. CORRELATION-BASED SKEW CALIBRATION

As shown in [4], the relationship between SNR and timing skew is a function of the input signal autocorrelation. This result can be rewritten in terms of the timing skews that maximize the SNR, i.e.

$$\arg(\max SNR) = \arg\left(\sum_{i=0}^{N-1} \max R(\tau_i)\right) \qquad (4)$$

Thus, the SNR is maximized when the autocorrelation of the individual sub-ADCs is maximized. This maximum is achieved with $\tau_i = 0$ for all $i = 0, ..., N-1$. In the scheme of Fig. 1, the autocorrelation is replaced by the cross correlation between each sub-ADC and the calibration ADC.

The interaction between the calibration ADC and each sub-ADC is further explained by focusing on just one sub-ADC, as in Fig. 4. The digital processor estimates the cross correlation, $\hat{R}(\tau)$, by averaging the product of the sub-ADC and calibration ADC output over M samples. Ignoring quantization effects, we have

$$\hat{R}(\tau) = \frac{1}{M}\sum_{n=1}^{M} y[n]y_c[n] = R(\tau) + E(M) \quad (5)$$

where $E(M)$ is the error term between the approximation $\hat{R}(\tau)$ and the actual cross correlation $R(\tau)$. Assuming stochastic signals, the variance of $E(M)$ is inversely proportional to M [8]. The value of the cross correlation is a function of the timing difference between the sampling edges of the two ADCs, and is maximized when the timing difference is reduced to zero.

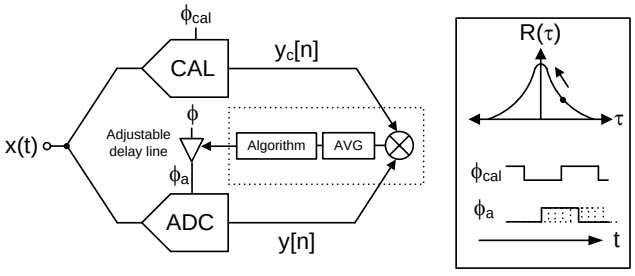

Fig. 4. Timing adjustment to maximize the correlation between a sub-ADC and the calibration ADC.

The overall objective of the algorithm block of Fig. 4 is to maximize the value of the cross correlation by climbing up the correlation curve. This is achieved by adjusting the delay line such that the sampling edge of the sub-ADC is adjusted in a direction that minimizes the timing difference and thus maximizes the correlation, as indicated in Fig. 4. Both LMS-based algorithms, in which the gradient of the cross correlation is estimated, and direct search algorithms can be used to adjust the control words of the delay line. These are implemented in the algorithm block of Fig. 4 as described in more detail in [9]. Although gain and offset errors do not affect the convergence of the algorithm, bandwidth mismatch introduces phase-shifts that can be problematic if not managed by proper circuit design considerations or additional correction blocks.

Since the cross correlation between the sub-ADC and the calibration ADC does not require the transfer function of both ADCs to be identical, it is possible to reduce the resolution of the calibration ADC to a single bit, as done in [2]. Reducing the resolution of the calibration ADC does not change the shape of the correlation function. If $R(\tau)$ is the autocorrelation of a signal $x(t)$, the correlation between $x(t)$ and some nonlinear function of $x(t)$ is simply a scaled version of $R(\tau)$ [10], which implies that the resolution of the calibration ADC can be reduced without loss of detail in the correlation function.

The main condition for the algorithm to work is that the correlation must change with the delay line codes; otherwise, the algorithm will not detect a global maximum and thus will not converge. The ADC quantization can limit the types of applicable signals; for example, a sinusoidal input signal with a frequency of f_s would prevent convergence. However, with sufficiently stochastic signals, as present in many communication systems, quantization is not a problem for proper detection of the cross correlation. A treatment on the required conditions for the input signal is provided in the next section.

4. CONDITIONS ON THE INPUT SIGNAL

In order for the calibration algorithm to work, the input signal $x(t)$ must have signal activity around the calibration ADC's decision levels. In case of a single-bit calibration ADC, there is a single trip point at $x = 0$. Furthermore, some stationarity conditions on the input signal are required, since the calibration algorithm estimates the correlation over a period of time and compares it to previous correlation values. Since all that is required for the algorithm is the value of the correlation, wide-sense stationarity requirements are sufficient. This can be relaxed if the signal is sample-invariant, such that

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=m_1}^{N+m_1} x(nT_s)\cdot x(nT_s-\tau) = \lim_{N\to\infty} \frac{1}{N}\sum_{n=m_2}^{N+m_2} x(nT_s)\cdot x(nT_s-\tau) \quad (6)$$

where $x(t)$ is the input signal and $m_1 \neq m_2$. However, it can still be acceptable if (6) is not true, as long as the autocorrelation changes "slowly" compared to the calibration algorithm convergence speed.

The other condition on the input signal is that the autocorrelation must have a single maximum within the region of concern. This is defined as the expected skew the ADC suffers from. For example, if the sub-ADCs suffer from timing skew of at most ±20 ps, then the region of concern is 20 ps. If there exist more than one maximum in this region, then there is no guarantee that the algorithm will converge to the correct value of τ. A sufficient condition to ensure this is convexity of the correlation function.

For application specific ADCs in which the signal autocorrelation function is known, the region of concern can be determined. However, generic bounds are useful, and can be derived by relating the autocorrelation to the signal power spectral density [11]. Assuming a differentiable autocorrelation and a real power spectral density, we have

$$R(\tau) = \int_{-\infty}^{\infty} G(f) e^{j2\pi f\tau} df \quad (7)$$

Taking the derivative results in

$$\frac{dR(\tau)}{d\tau} = 2\pi \mathrm{j} \int_{-\infty}^{\infty} fG(f) e^{j2\pi f\tau} df = -4\pi \mathrm{j} \int_{0}^{\infty} fG(f)\sin(\pi f\tau)\, df \quad (8)$$

The region of concern is derived by having

$$\frac{dR(\tau)}{d\tau} \leq 0 \quad (9)$$

for all τ in $-\tau_{max} \leq \tau \leq \tau_{max}$. Thus, τ_{max} defines the region that guarantees a single maximum. For example, if $G(f)$ is band limited to B such that $G(f) = 0$ for $f > B$, then $\tau_B = 1/2B < \tau_{max}$, as seen from (8). In other words, if the expected timing skew is ±20 ps, then an input signal band limited to 25 GHz is guaranteed to have a single maximum in this region. Although sufficient, such a condition is not necessary. An input signal with a first-order low pass power spectral density is not band limited, but has an autocorrelation function that is monotonically decreasing for $\tau > 0$, and thus has a single maximum for all τ.

Finally, we consider in this section the effect of quantization, which was ignored in the preceding analysis. The correlation was calculated through an approximation, and in (5), the variance of $E[M]$ only falls off with $1/M$ when the samples of the errors are uncorrelated [8]. This is not the case with quantization noise.

As a trivial example to illustrate this, assume an input signal of $x(t) = \cos(2\pi f_{in} t)$, where $f_{in} = f_s$. The output of the single-bit calibration ADC is $y_c[n] = \mathrm{sign}[\cos(2\pi n)] = 1$ for all n. If the sub-ADC also has single bit resolution, then its output is $y[n] = \mathrm{sign}[\cos(2\pi(n - f_{in}\tau))]$, which is equivalent to

$$y[n] = \begin{cases} 1 & if\ \frac{1}{4f_{in}} \geq \tau \geq \frac{1}{-4f_{in}} \\ 0 & else \end{cases} \quad (10)$$

Thus, the value of the correlation does not change as long as $1/4f_{in} \geq \tau \geq -1/4f_{in}$, which means that the timing skew cannot be corrected because of the sub-ADC quantization. Note that this is not the case if the sub-ADC has infinite resolution such that $y[n] = \cos(2\pi(n - f_{in}\tau))$.

Therefore, quantization errors in the sub-ADCs can be problematic for the algorithm. However, assuming stochastic signals, if the input signal $x(t)$ is stationary, ergodic, continuous, and has a non-zero probability of crossing $x = 0$, then there is a non-zero probability that a zero-crossing exists between $nT_s - \tau$ and nT_s for all τ [12]. In other words, if enough samples are collected, then the number of zero-crossings decreases with τ, which is sufficient to ensure that the correlation function will increase and not suffer from the effects of quantization.

This is not guaranteed in sinusoidal signals, as already illustrated. If the ratio of the input frequency to the sampling frequency is irrational, then collecting enough samples will guarantee that zero-crossings exists between $nT_s - \tau$ for all τ, since there will exist samples where the value at nT_s and $nT_s - \tau$ do not have the same sign. If the ratio is rational, such that $f_{in}/f_s = N/M$ for integers N and M where the greatest common divisor of N and M is 1, then this is not guaranteed since the samples are periodic in M. However, if $1/M \leq \Delta\tau$, where $\Delta\tau$ is the delay line step size, then although the zero-crossings may not monotonically decrease as τ is continuously adjusted, they will decrease as τ is discretely adjusted with the delay line step size. As the resolution of the sub-ADC increases, this becomes less of a problem, and the number of frequencies for which the calibration algorithm will not properly work decreases.

5. RESULTS

The discussed calibration was implemented and tested using an experimental 5-bit, 12-GS/s time interleaved ADC using 8 interleaved channels and a single-bit calibration ADC [2]. The algorithm was tested using sinusoidal inputs of various frequencies and amplitudes. For the results shown in Fig. 5, a sinusoid with a frequency of 8 GHz was applied at the input and the calibration ADC was clocked at 480 MHz. The ADC output was decimated by a factor of 81 to simplify the test hardware. The delay line values were controlled by a standard gradient based stochastic maximizer, which is described in more detail in [9].

As shown in Fig. 5(a), the signal-to-noise and distortion ration (SNDR) improved from approximately 12 dB to around 24 dB once the calibration is turned on, and converged to a stable point within 20 calibration cycles. In this example, each calibration cycle consisted of 500,000 samples, which requires approximately 8 ms. This results in a total start up time of approximately 160 ms. Fig. 5(b) shows the timing skew calibration codes that are fed to the delay lines and which are updated at the end of each calibration cycle.

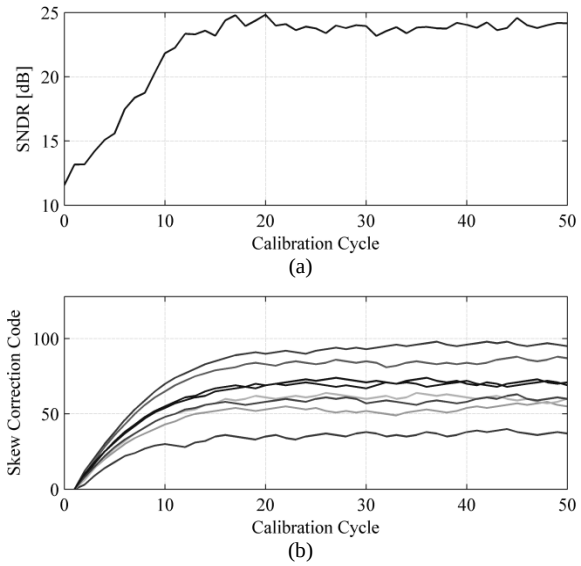

Fig. 5. Measured timing skew calibration data. (a) SNDR convergence and (b) timing skew correction codes.

6. REFERENCES

[1] W. Black and D. Hodges, "Time interleaved converter arrays," *IEEE J. Solid-State Circuits*, vol. 15, no. 6, pp. 1022-1029, Dec. 1980.

[2] M. El-Chammas and B. Murmann, "A 12-GS/s 81-mW 5-bit time-interleaved flash ADC with background timing skew calibration," in *Symp. VLSI Circuits Dig.*, Jun. 2010, pp. 157-158.

[3] K. Poulton et al., "A 20 GS/s 8 b ADC with a 1 MB memory in 0.18 μm CMOS," in *ISSCC Dig. Tech. Papers*, Feb. 2003, pp. 318-496.

[4] M. El-Chammas and B. Murmann, "General Analysis on the Impact of Phase-Skew in Time-Interleaved ADCs," *IEEE Trans. Circuits Syst. I*, vol. 56, no. 5, pp. 902-910, May 2009.

[5] S. 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.

[6] C. Vogel and H. Johansson, "Time-interleaved analog-to-digital converters: status and future directions," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2006, pp. 3386-3389.

[7] A. Haftbaradaran and K. Martin, "A Background Sample-Time Error Calibration Technique Using Random Data for Wide-Band High-Resolution Time-Interleaved ADCs," *IEEE Trans. Circuits Syst. II*, vol. 55, no. 3, pp. 234-238, Mar. 2008.

[8] W. Root and W. Davenport, *An Introduction to the Theory of Random Signals*, McGraw-Hill, 1958.

[9] M. El-Chammas, "Background Calibration of Timing Skew in Time-Interleaved A/D Converters," Ph.D. dissertation, Stanford University, 2010.

[10] J.J. Bussgang, "Crosscorrelation functions of amplitude-distorted gaussian signals," *MIT Research Laboratory of Electronics Technical Reports*, no. 216, Mar. 1952.

[11] F. Todero, "On some bonds between autocorrelation and power spectra functions," *Proceedings of the IEEE*, vol. 56, no. 12, pp. 2170-2171, Dec. 1968.

[12] J. McFadden, "The axis-crossing intervals of random functions," *IRE Transactions on Information Theory*, vol. 2, no. 4, pp. 146-150, Dec. 1956.