A 10-b 800MS/s Time-Interleaved SAR ADC with Fast Timing-Skew Calibration

Jeonggoo Song, Kareem Ragab, Xiyuan Tang, and Nan Sun
University of Texas at Austin, Austin, TX, USA
Email: Jeonggoo.song@utexas.edu, nansun@mail.utexas.edu

Abstract—This paper presents a time-interleaved (TI) successive-approximation-register (SAR) ADC with a fast variance-based timing-skew calibration technique. It uses a single comparator-based window detector to calibrate the timing skew. It has low-hardware cost and $10^4$ times faster convergence speed compared to prior variance-based timing skew calibration technique. The proposed technique brings collateral benefits of offset mismatch calibration. A prototype 10-b 800MS/s ADC in 40nm CMOS achieves Nyquist-rate SNDR of 48 dB and consumes 9.8mW, leading to a Walden FoM of 59-IF/conversion-step.

Keywords – time-interleaved ADC, variance-based timing-skew calibration

I. INTRODUCTION

TI SAR ADC is well-known for its energy-efficiency for high-speed and medium-resolution applications. However, without calibrating mismatches (e.g. gain error, offset error, and timing skew), it is nontrivial to achieve good linearity. Among all, timing skew mismatch, which aggravates at high input frequencies, is known to be a linearity bottleneck. Various timing skew calibration techniques have been developed. In [1], a dedicated reference ADC was used. The reference ADC is a replica of the TI ADC channels and runs at $f_s/(N+1)$, where $N$ is the number of TI channels. Having a full-blown ADC replica increases power, especially for a small $N$. Additionally, it sets a lower limit on the alignment (beat) period of the reference and each ADC channel to $N(N+1)$ clock cycles, reducing the calibration speed. Furthermore, the alternatingly operated reference channel introduces spurs by periodically changing ADC input impedance. The reference channel is simplified to a single comparator which reduces power [2]. However, its convergence speed is slow for random inputs due to large autocorrelation estimation errors. Flash-assisted time-interleaved (FATTI) SAR addresses these problems by having a low-resolution flash ADC running at the ADC full rate [3]. Timing skew is minimized by reducing the variance of the difference between each channel’s output and the flash ADC. However, the flash ADC consumes large power. Moreover, the convergence speed is still slow when operating in the background with random inputs that cause large fluctuations in the variance estimation.

This paper presents a novel variance-based timing-skew calibration technique for TI ADC. It exploits the relationship between the comparator input and its decision time to identify input samples that are close to the comparator threshold. By using only these samples, the variance computation has much less fluctuation. As a result, the proposed technique substantially reduces the number of samples needed to obtain an accurate variance estimation which consequently significantly boosts the convergence speed. Our prototype ADC requires only $10^5$ input samples per calibration step, which is 4 orders of magnitude smaller than that of the prior variance-based timing-skew calibration technique of [3]. Furthermore, the proposed technique also obviates the need for a flash ADC; instead, it only requires a single comparator-based window detector, which reduces hardware overhead and power.

This paper is organized as follows. Section II introduces the proposed timing-skew calibration technique. Section III presents the convergence time analysis. Measurement results are discussed in Section IV. Conclusion is brought up in Section V.

II. PROPOSED TIMING-SKEW CALIBRATION TECHNIQUE

A. Proposed Time-Interleaved ADC Architecture

Fig. 1 shows the architecture of a 4-way TI SAR with the proposed timing-skew calibration technique. It consists of 4 SAR ADCs, a window detector based reference channel, an output MUX, a multi-phase clock generator, variable delay lines (VDLs), and a timing-skew calibration engine. The window detector operates at full ADC rate, so that the ADC input impedance is constant at every clock cycle, eliminating potential spurs. The calibration engine uses the window detector to check whether a sampled input $V_{in}$ is within a fixed window, i.e., $|V_{in}| \leq W$, where $W$ is the window width. This window detector consists of a replica sampling network, a dynamic latch comparator, a current-starved inverter chain, and a flag DFF. The window flag $f$ is raised when the clock signal arrives at the DFF before the XOR output goes high. The inverter chain is configured so that $\tau_{delay} = \tau_{comp, LSB} + \tau_{XLOR}$, where $\tau_{XLOR}$ is the XOR delay and $\tau_{comp, LSB}$ is the comparator delay when $|V_{in}| = W$. If $|V_{in}| \leq W$, regeneration time of the comparator is longer than $\tau_{comp, LSB}$ and the flag $f$ is raised ($f = 1$). On the other hand, if $|V_{in}| > W$, the flag $f$ is not raised ($f = 0$). $W$ can be tuned by adjusting $\tau_{delay}$. Increasing $\tau_{delay}$ reduces $W$, while reducing $\tau_{delay}$ enlarges $W$. $\tau_{delay}$ is adjusted by changing the bias current into the current-starved inverter chain of the window detector.

B. Proposed Timing-skew Calibration Technique

The timing-skew calibration engine collects ADC outputs with $f = 1$ and sorts them into 4 sets depending on which ADC channel they come from. Then, the engine calculates the variance for each set of channel outputs. With a large timing skew, the accumulated channel output samples are scattered to side-corners, as shown in the Fig. 2 for a sinusoidal input, which leads to a large variance. By contrast, without any timing skew, the accumulated ADC outputs with $f = 1$ are distributed among few codes and the variance is minimum. Thus, the variance is an indicator for the timing skew. Timing skew is reduced as variance is decreased. Therefore, timing skew is removed by...
minimizing the variance. The VDL is controlled by the calibration engine to align the sampling edges of the reference channel and 4 ADCs. VDL delays are added or removed in the direction achieving lower variance. The digitally controlled VDL of this work has 5b binary-weighted control codes for fine delay tuning of 300fs per step and 2b coarse control codes with 2ps per step.

III. CONVERGENCE TIME ANALYSIS

Convergence time is an important factor for any timing-skew calibration algorithm. Convergence time must be short enough to keep track of temperature changes because the delay is highly sensitive to temperature. For the proposed technique, convergence time is limited by the number of input samples, $M$, that is required to obtain an accurate estimation of the variance.

The key parameter that sets $M$ is the window width $W$. At first glance, it seems that a large $W$ is preferred as it increases the number of ADC inputs falling into the window, and thus, can reduce error in variance estimation by averaging. Nevertheless, a large $W$ greatly increases the variance fluctuation for ADC samples with $f=1$. This significantly increase the convergence time, as a much bigger number of input samples should be averaged in order to obtain an accurate estimation of the variance.

To prove this, let us first calculate the variance of an ADC channel's outputs with $f=1$, denoted as $\sigma^2_2(\tau)$:

$$\sigma^2_2(\tau) \equiv \text{Var}(V_{in}(t + \tau)|f=1)$$
$$\approx \text{Var}(V_{in}(t)|f=1) + \tau^2 \cdot \text{Var}\left(\frac{dv_{in}}{dt}\right)$$ (1)

where $\tau$ is the timing skew between the ADC channel and the window detector. Here, in deriving (1), we have exploited the fact that $V_{in}(t)$ and $dv_{in}/dt$ can be treated as two uncorrelated random variables. The $2^{nd}$ term on the right hand side of (1) carries information on $\tau$ and is the “signal” component in the time skew calibration process. By contrast, the $1^{st}$ term on the right side of (1) contributes only “noise” in the overall variance estimation. We denote the RMS of the estimation error of (1) term as $N_{est}(\tau)$. It can be derived that [4]:

$$N_{est}(\tau) \approx \frac{1}{M \cdot p(f = 1)} \cdot \text{Var}(V_{in}(t + \tau)|f=1)$$
$$\approx \frac{1}{M \cdot p_0 \cdot W} \cdot \text{Var}\left((V_{in} + \tau \frac{dv_{in}}{dt})|f=1\right)$$
$$= \frac{1}{M \cdot p_0 \cdot W} \cdot \left(45 W^4 + \frac{4}{3} W^2 \tau^2 \sigma_{av}^2 + 2 \tau^4 \sigma_{dv}^4\right)$$ (2)
where \( P(f=1) \) is the probability of an input sample \( V_{in} \) falling inside the window, and it is given by the product of the window width \( W \) and the probability density of \( V_{in} \) inside the window, denoted as \( p_0 \). In deriving (2), we have made the following approximations: i) \( V_{in} \) is uniformly distributed within the window \([-W, +W] \); ii) \( dV_{in}/dt \) follows normal distribution with the standard deviation of \( \sigma_{v0} \). As shown in (2), although \( P(f=1) \) increases linearly with \( W \), the fluctuation in the variance estimation, which is captured by \( Var(V_{in}(t+\tau)f_{in}=1) \), increases with \( W^4 \), which greatly outweighs the benefits of increased \( P(f=1) \). For a given timing-skew calibration accuracy, \( N_{est}(\tau) \) is fixed, which implies that \( M \) is proportional to \( W^4 \). This result has been confirmed via behavioral simulations, as shown in Fig. 3. Here, the input \( V_{in} \) is assumed to be a Gaussian random input with bandwidth of \([0, f_s/4]\) and \( \sigma = 0.2V_{ref} \). This input model is suitable for many practical applications, e.g., full-band capture receivers. In summary, to accelerate convergence speed, a small \( W \) is preferred.

In the prototype ADC, \( W \) is set to 1 LSB, limited by the quantization step. Under this setup, simulation results show that the proposed calibration technique requires about \( 10^6 \) samples per channel for each variance estimation step in order to reach 1ps calibration accuracy. With each ADC channel operating at 200MS/s, this translates to a short calibration cycle of only 0.5ms. By contrast, under the same condition, the calibration technique of [3] requires about \( 10^9 \) samples, which is 1000 times slower. The reason for this much longer convergence time is that a 4-bit flash is used in [3], and thus, its equivalent window size is 64 LSB, leading to a much larger error in the variance estimation, which requires significantly larger number of samples for averaging. Additionally, the proposed calibration technique requires only a small fraction of the \( 10^6 \) samples (only those falling inside the window) for variance computation. This further simplifies its implementation and reduces its hardware cost and power.

As in any timing skew calibration scheme, the proposed technique has requirements on the input signal. It does not work for inputs spanning only a few non-zero LSB bins (e.g., DC signals). It requires the input to have samples that fall into the window. Such requirement, however, is not difficult to satisfy in practical situations. For example, the proposed technique works well for both narrow-band and wide-band random signals. It also works for pure sinusoidal signals, as long as its frequency is not exactly a fraction of \( f_s \) (e.g., \( f_{in}=f_s/10 \)).

The proposed calibration technique can also be used to correct offset mismatches. Channel offsets can be extracted from mean values of ADC outputs with \( f = 1 \). The offset is digitally cancelled by subtraction from each channel output.

### IV. MEASUREMENT RESULTS

The proposed calibration technique is applied to an 800MS/s 4-way TI ADC built in 40nm CMOS. Each channel is a 10-b 200MS/s asynchronous SAR ADC. Both timing-skew and offset mismatches are removed by the proposed calibration technique. Gain mismatch is not observed due to good capacitor matching. After timing skew calibration, the variance of each channel reaches to its minimum as shown in Fig. 4. Fig. 5 demonstrates the measured distributions of variance samples of a single ADC channel with \( t_{skew} = 0, 1 \text{ps}, \text{and } 2 \text{ps} \). Each distribution consists of 150 samples and the measurement is done with \( f_{in} \approx f_s/2 \) and \( W \approx 1 \text{ LSB} \). In Fig. 5, \( \mu \) denotes the mean of sampled \( \sigma_v^2 \) of equation (1) and \( N_{est} \) represents the standard deviation of the distribution. The measured data of \( \mu \) verifies the relationship between \( \sigma_v^2 \) and \( t_{skew} \) of (1). From (1), \( \sigma_v^2 \) of \( t_{skew}=2ps \) is expected to be approximately \( (2\text{ps})^2 = 4 \text{ times} \) greater than \( \sigma_v^2 \) of \( t_{skew}=1ps \). As expected, the measurement results verify that \( \mu_{t_{skew}=2ps} \) is almost 4 times greater than \( \mu_{t_{skew}=1ps} \cdot N_{est} \), which is also given by (2), indicates how much the distribution fluctuates from its mean. The relationship between \( N_{est} \) and timing skew of (2) is verified by comparing the expected result from (2) and measurement data of Fig. 5. In order to estimate \( N_{est} \) for \( t_{skew}=2ps \), each term in (2) were deduced from data of \( N_{est} \) when \( t_{skew}=0 \) and 1ps. Using the estimated parameters, the estimated \( N_{est}|t_{skew}=2ps \) using (2) is 1.14 which matches the measured \( N_{est} \) of 1.3. The relationship between \( N_{est} \) and \( M \) in (2) is also verified from measurement results in Fig. 5. From (2), \( N_{est} \) is expected to increase by \( \sqrt{10} \) when \( M \) is reduced by 10 times. The measured \( N_{est} \) of \( t_{skew}=1ps \) is about 3 times larger when \( M \) is reduced tenfold.
Moreover, Fig. 5 measurement results confirm the simulation results in Fig. 3, which indicates that $M = 10^5$ samples are required in order to suppress fluctuations in overall variance estimation with 1ps calibration accuracy. Fig. 5a demonstrates three distributions when each variance sample is collected with $M = 10^5$. Three distributions are far apart from each other and their fluctuations ($N_{est}$) are small enough that no overlapping occurs among them. Fig. 5b observes more closely on two variance distributions with $t_{\text{skew}} = 0$ and 1ps, when $M$ is reduced to $10^4$. $N_{est}$ of both variance distributions increase. The variance estimations fluctuate substantially and their distribution overlap. Therefore, when $M$ is reduced to $10^4$, the calibration engine will have higher chance of controlling VDL in wrong direction, due to an error in variance estimation. Fig. 6 shows the measured ADC output spectrum before and after skew calibration. When the timing-skew calibration reduces the variances of 4 channels to their minimums, tones from timing skews are reduced by more than 10 dB. SNDR of 48dB and SFDR of 56dB are achieved for Nyquist rate inputs. Fig. 7 shows measured INL and DNL after the timing-skew calibration. The INL and DNL are $+1/-1.5$ LSB and $+0.7/-0.6$ LSB, respectively. Fig. 8 shows the chip microphotograph and the active area is 500 $\mu$m by 300 $\mu$m . Clock generator is located in the center and the 4 SAR ADCs are symmetrically around the clock generator. The window detector is located in the right corner of the clock generator and it only occupies relatively small area compared to the other channels. The total ADC power is 9.8mW with 1.1V power supply. The ADC achieves Walden FoM of 59fJ/conv-step at Nyquist input and it is comparable to the state-of-the-art. Table I summarizes ADC performance.

V. CONCLUSION

This paper presents a 4-way TI SAR ADC with a novel variance-based timing-skew calibration technique. It can calibrate timing skew as well as offset mismatches among all channels. It achieves a much faster convergence speed than the prior variance-based calibration technique. It also reduces the hardware cost by replacing a flash ADC by a single-comparator based window detector.

ACKNOWLEDGEMENTS

The authors thank TSMC for the chip fabrication. This work is supported in part by NSF grants 1254459, 1509767, and 1527320.

REFERENCES


