IEEE Open Journal of the
Solid-State Circuits Society

Received 6 July 2022; revised 11 September 2022 and 30 September 2022; accepted 30 September 2022. Date of publication 5 October 2022; date of current version 24 October 2022.

Digital Object Identifier 10.1109/OJSSCS.2022.3212028

An Introduction to High Sample Rate Nyquist Analog-to-Digital Converters

GABRIELE MANGANARO (Fellow, IEEE)

(Invited Paper)

Strategic Technology Exploration Platform, MediaTek Inc., Woburn, MA 01801, USA

CORRESPONDING AUTHOR: G. MANGANARO (e-mail: gabriele.manganaro@ieee.org)

ABSTRACT Increasingly wider band analog signals found in multiple information and communication technology applications, requiring real-time digital signal processing, demand analog-to-digital converters with ever higher sample rate. Several innovative techniques, from the circuit level, to architecture and algorithms, have enabled remarkable breakthroughs in a relatively short span of time. This overview article aims to introduce this topic and to point to some of the most notable results, while also highlighting open problems and engineering trends.

INDEX TERMS Analog circuits, analog–digital conversion, digital–analog conversion, digitally assisted analog (DAA) circuits, S/H, T/H.

I. INTRODUCTION

NUMEROUS applications in areas, such as communication, instrumentation, and sensing, require the acquisition and processing of analog signals with increasingly wider bandwidth and, hence, demand the use of ever higher sample rate analog-to-digital converters (ADCs) [1], [2], [3], [4] in their systems.

Over the past few decades, transistors' speed increased when migrating to finer lithography nodes due to CMOS process feature scaling ("Dennard's scaling") [5]. Although analog designers dealt with nontrivial challenges due to lowering power supplies, charge leakage, device noise, and so on, speed and power consumption efficiency benefitted from Dennard's scaling [6].

Nowadays, feature scaling in newer MOS processes continues enabling digital circuits with an ever smaller area, cost, and power consumption per function [7], [8] but the same cannot be said for analog circuits. In the case of analog circuits, the transition from planar MOSFETs to FinFETs processes improved key transistor parameters and provided higher drain impedance, reduced gate and subthreshold leakage, greater transconductance efficiency, device matching, etc. [9], [10]. But generational improvements in intrinsic device speed have substantially diminished [11] while interconnect parasitics, particularly resistive ones, have worsened, limiting the circuits' high-frequency performance [12].

Despite that, ADCs' sample rate f_s breakthroughs continue unabated, as shown in Fig. 1, reporting some of the fastest reported ADCs [13], for three classes of signal-to-noise and distortion ratio (SNDR). Fig. 1 also shows the transition frequency f_t, as a proxy for process' speed, at the time in which these processes were first utilized in these data converters [14].

Over the last decade, high sample rate ADCs have got 10–100 times faster, despite transistors not getting much faster. New process technologies do support data converters speed progression, though not in the same manner as in the past. As shown in this article, the explanation for the dramatic f_s speed-up shown in Fig. 1 can be found in a combination of the circuit, architectural, and algorithmic innovation.

In the following, it is assumed that the reader is familiar with the general concepts of ADC design. The objective of this article is to introduce the specifics of Nyquist-rate high-speed ADCs, provide a high-level overview of the most recent and notable results, and highlight some of the open problems and active lines of research and some possible future directions in a field that is vast and rapidly evolving. Throughout this article, developing intuition is intentionally

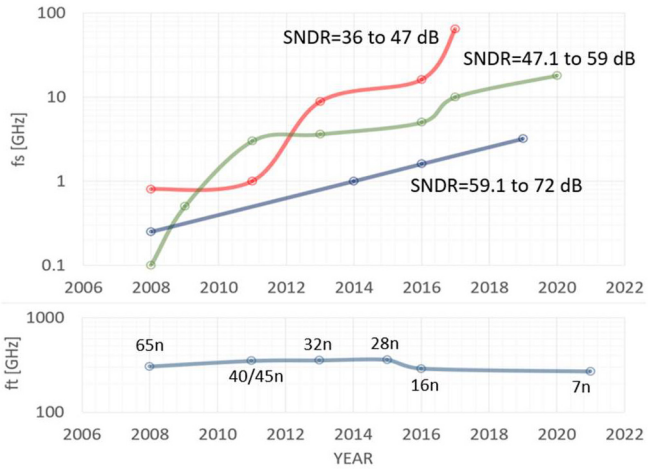

FIGURE 1. Highest ADCs' f_s and process nodes' f_t published year on year at major conferences [13], [14].

prioritized, providing high-level descriptions, while deferring the reader to the bibliography for in-depth study.

The remainder of this article is organized as follows. In Section II some background material is summarized, to help frame some of the differences in objectives and corresponding primary metrics between high-speed ADCs and general-purpose ADCs. In Section III, two core speed-up techniques are described, highlighting their broad applicability. In Section IV, the time-interleaving (TI) of ADCs is discussed along with its advantages and the nonidealities that it introduces. The calibration of TI ADCs is the topic of Section V. All the concepts and ideas covered in the previous section come together in Section V, where targeted examples of recent high-speed ADCs are provided. Conclusive remarks in Section VI complete this article.

II. BACKGROUND

At the time of writing of this article, the term *high sample rate* ADC is used for an ADC with $f_s > 1$ GHz (GSPS). *Medium resolution* refers to resolution n between 8 and 14 bits. The term *high-speed ADCs* is used to denote high sample rate *and* medium-resolution ADCs [15], [16].

High-speed ADCs have multiple applications but are not for general purposes. Therefore, before discussing implementation specifics, it is necessary to clarify the specific design objectives and their metrics, some of which are *different* than those used for general-purpose ADCs as discussed in Section II-A. Moreover, independently of the quantizer architecture, one of the most prominent limitations for high-frequency signal capture is the sampling aperture error, briefly discussed in Section II-B.

A. MAIN METRICS

In this section, the reasons for using RF design metrics, such as noise spectral density (NSD) or intermodulation distortion (IMD), is explained in the context of high-speed ADC's

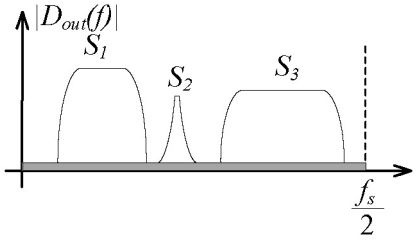

FIGURE 2. Multiple signals composing a wide-band ADC's digitized output.

performance. The interested reader is referred to the literature for the formal definitions [15], [16], [17], [18].

High-speed ADCs' outputs have specific features. Their (zero-input) digitized noise is frequency-uniform (white) and Gaussian. The ADC's nonlinearity, ignoring quantization, is weakly nonlinear, mainly polynomial [17], [18].

Let us denote the ADC's continuous-time analog input signal with v_{in} and the discrete-time digital output with D_{out}, respectively. The digitized noise in D_{out} is dominated by thermal noise. The power associated with the quantization nonlinear distortion (often referred to as *quantization noise*) is negligible compared to the thermal noise power, by design. Hence, quantization and resolution n are not insightful metrics for this class of ADCs.

In applications, such as in wireless communications, high-frequency instrumentation, and instances of wide-band sensing, the output spectrum is often composed of the aggregation of multiple band-limited signals, such as in the example shown in Fig. 2.

Some of these signals may be desired and, after conversion, are digitally separated and processed (e.g., S_1 and S_3), while others may be jammers (e.g., S_2) to be digitally filtered. During the system operation, part of these co-existing signals can disappear, leaving portions of the spectrum temporarily open, while new ones emerge, occupying available frequency bands. That is descriptive, for example, of the operation of a high-speed ADC employed in a Wi-Fi access point, an IoT gateway, or a wireless base-station radio head, where different client devices connect to/disconnect from the gateway, in the presence of jammers [17]. Hence, when considering $D_{\text{out}}(f)$, attention is paid to the ratio between the power of the desired signal and the power of the noise (or of undesired signals) at adjacent frequencies. NSD, measured in the frequency range close to the desired signal's band edge, becomes a more important metric than the overall signal-to-noise ratio (SNR) evaluated over the entire first Nyquist range $[0, f_s/2]$. For instance, referring to the spectrum shown in Fig. 2, the NSD in the frequency range between S_1 and S_2 and between S_2 and S_3 is a key specification since it directly relates to the system receiver sensitivity in such a use case [17].

Furthermore, one of the roadmap objectives for high-speed ADCs is to extend f_s to a higher frequency, one generation after the another, while keeping NSD ideally constant. For instance, in Fig. 2, extending f_s allows digitizing one, or

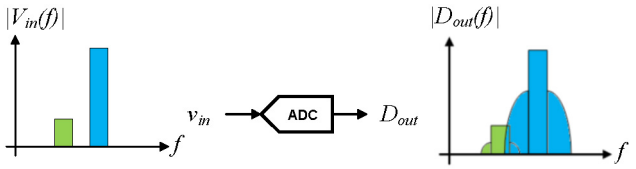

FIGURE 3. Desensitization of a weak band adjacent to a strong band due to IMD.

more, higher frequency signals (to the right of S_3). This generational progression, however, corresponds to decreasing values of SNR (due to the integration of the fixed NSD over larger frequency intervals $f_s/2$) and makes the SNR less representative of the objective.[1]

Quantifying nonlinear distortion is also important, particularly in the case of two or more adjacent signal bands in $D_{\text{out}}(f)$. The ADC's nonlinearity causes mutual interference between these distinct signals: spurious intermodulation of v_{in} is introduced in D_{out} [17], [18]. When a weaker (lower power) band/channel is adjacent to a stronger (higher power) band/channel, as for $V_{\text{in}}(f)$ in Fig. 3, IMD between the two adjacent bands generates spurious signal power in the adjacent frequency range: resulting in undesired spectral power tails added on $D_{\text{out}}(f)$ to the sides of the two input bands. The spurious power originated by the stronger signal band, encroaching the weaker signal, is concerning: this interferes and degrades the weaker signal's quality, possibly making it undetectable (this is referred to as *desensitization*) [14], [15].

Nonlinear distortion is quantified by metrics, such as the third-order IMD (IMD3), fifth-order IMD (IMD5), and higher [18], or by the adjacent channel leakage ratio (ACLR) [15], [16], [17], [18]. Rarely, it is quantified by total harmonic distortion (THD), which is of general-purpose use. The spurious-free dynamic range (SFDR) is also a key metric of spectral purity [15], [17]. SNDR and effective number of bits (ENOB), which lump all spurious content regardless of its manifestation, are less insightful metrics.

For example, let us consider the three ADCs in [20], [21], and [22] with main metrics reported in Table 1. From an application point of view, the ADC in [20] is superior to the one in [21], which, in turn, is superior to the one in [22]. That is because of the increasingly higher sample rate while meeting the same NSD. But the SNDR (and ENOB) erroneously suggests the opposite. Also, while sufficiently high-resolution n is required to ensure that the total noise is white and Gaussian, n is not indicative of the conversion precision as confirmed, for instance, by its gap with ENOB.

In conclusion, for high-speed ADCs, greater attention is paid to NSD than SNR; IMD or SFDR are more relevant than THD, and resolution n is not a primary concern if it is at least two bits higher than ENOB.

1. Other applications include RADAR, LIDAR, time-of-flight applications etc., where chirps or other complex input signals also demand for increasingly higher sample rate [19]. For these, time-domain specifications for settling time and accuracy are used, making again SNR not directly insightful.

TABLE 1. Example of metric comparison between high-speed ADCs.

Symbol	Quantity	[20]	[21]	[22]
f_s	Sample rate (GSPS)	18	10	5
n	Resolution (bits)	12	12	14
P	Power consumption (W)	1.3	2.9	2.3
NSD	Noise spectral density (dBFS/Hz)	-157	-157	-157
SNDR	Signal-to-Noise and Dist. Ratio (dB)	52	55	58
ENOB	Effective number of bits (bits)	8.35	8.84	9.34

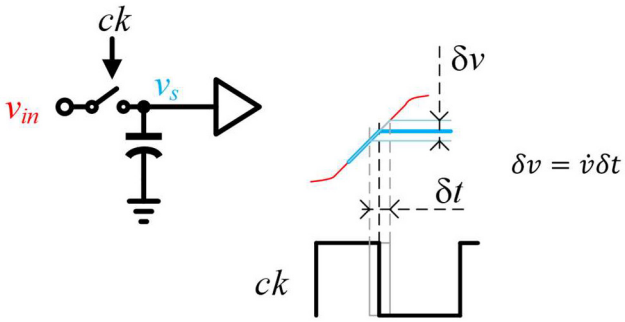

FIGURE 4. Sampling voltage error δv on v_s resulting from clock *ck* timing error δt [23].

B. SAMPLING CLOCK APERTURE ERROR

It is impossible to talk about high-speed ADCs without discussing the impact of clock jitter to the sample time error [23], [24]: a nonideality affecting the sampled voltage v_s of the analog input v_{in}, is the *aperture error* δv caused by time jitter δt in the sampling clock *ck*, as in the switched capacitor track-and-hold (THA) shown in Fig. 4.

Due to the clock *ck*'s phase noise, the sample time of v_{in}, determined by the falling edge of *ck*, is subject to a random time error δt, quantified by the rms jitter σ_j. Hence, v_s is affected by a random error δv, manifesting itself as additive voltage noise. This does not (directly) depend on f_s as it solely depends on what occurs at the time of aperture of the sampling switch, not by the entire sampling period or its duration. In addition to the time jitter σ_j, the noise introduced on v_s also depends on the rate of change of v_{in} at the time of sampling. In the simple case of a sinusoidal input $v_{\text{in}} = A_{\text{in}}\cos 2\pi f_{\text{in}} t$, the additional total noise power introduced by the aperture error is $P_j = 2\pi^2 A_{\text{in}}^2 f_{\text{in}}^2 \sigma_j^2$. This sets an upper bound to the SNR: $\text{SNR}_{j(\text{dB})} = -20\log_{10} 2\,\pi f_{\text{in}} \sigma_j$ [25].

In short: in equal jitter conditions, sampling higher frequency inputs introduces more noise. This has system-level ramifications going beyond the ADC, since the power spent to *generate* the sampling clock, for instance, by means of a VCO servo-ed in a PLL loop, is inversely proportional to the power of the jitter $P_{\text{VCO}} \propto 1/\sigma_j^2 \propto \text{SNR}_j^2$ [25].

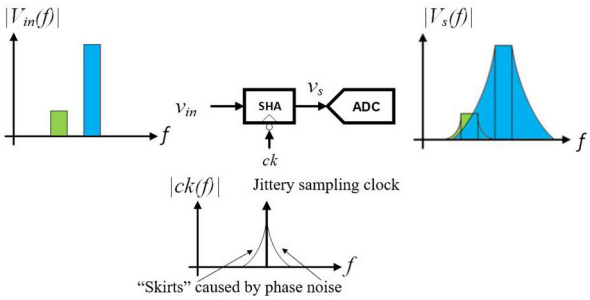

FIGURE 5. Desensitization caused by a jittery sampling clock.

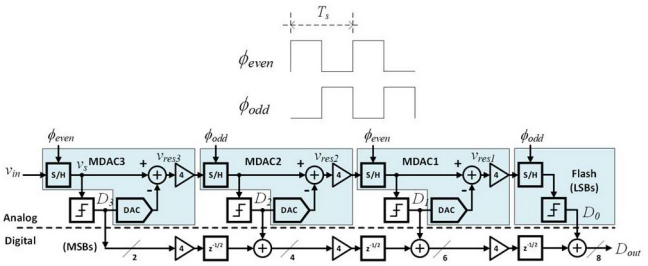

FIGURE 6. Pipelined ADC.

A closer look at it, however, highlights important points. First, similarly to the considerations made in the previous section, capturing the sampling time's uncertainty with a single total power number or its impact on the overall SNR degradation is not always insightful. The noise degradation that is of concern is the one in the range adjacent to the desired signal bands, not the total integrated error across the full Nyquist range. Because of the above-described mechanism, sampling aperture error specifically affects the higher frequency signals with the largest power. Not understanding this aspect results in unrealistic targets for σ_j and, hence, P_{VCO} [26].

The frequency profile of the phase noise is also very important as illustrated in Fig. 5 (in the case of two adjacent bands where the weak band risks desensitization) [14], [27].

The phase noise power of clock *ck* is frequency dependent. Its spectrum is white at large frequency offsets from the oscillation frequency then it gradually trends as $1/f$ for frequencies closer to the oscillation frequency (and as $1/f^{\alpha}$ with growing $\alpha > 1$ for smaller frequency offsets from the oscillation frequency). This can be seen in the spectrum of *ck(f)* shown in Fig. 5 for the case of a free-running oscillator.[2] This spurious noise power's shape is referred to as *noise skirts*. Sampling v_{in} with *ck* results into v_s "inheriting" the noise skirts of *ck* in an amount dependent on the rate of change of v_{in}. For high input frequency f_{in}, this spurious noise power can extend into an adjacent band. But it is the excess noise power leaking into the adjacent band edge that is of concern, rather than the total jitter σ_j. Finally, this nonideality occurs *before* and independently from the quantization: it is due to sampling, it is not due to the digitization.

While the impact of clock jitter is emphasized in this section, it is important to briefly mention that other nonidealities also contribute to the aperture error. For instance, if the sampling switch in Fig. 4 is implemented with a simple nMOS switch and its turn-on gate voltage is set by the clock's supply voltage, the sampling time will directly depend on v_{in} as this voltage determines the channel's potential (hence, the overdrive voltage) and, therefore, the actual instant in which the channel ceases to conduct [15]. Moreover, both the charge stored in the channel and in the device stray capacitances depends on v_{in} and as the gate drives the switch to cut off, the v_{in}-dependent discharge impacts the actual turn-off instant and also adds input-dependent charge to the capacitor, affecting the actual sampled v_s. Unlike the stochastic error caused by the clock's phase noise, these nonidealities introduce harmonic distortion to v_s. Multiple remedies exists for the v_{in}-dependence including constant-V_{gs} gate boosting [33], [34] and switch bootstrapping [15], [20], [21], [22], making the gate voltage track the channel voltage to keep the overdrive constant. Charge injection calibration [20], [22] has also been proposed as a linearization technique.

2. When the VCO is servo-ed in a PLL, the total phase noise of the synthesized clock is shaped by the loop's frequency response and by other noise sources at frequency offsets below the loop bandwidth [26]. Though, in this case, the higher offset range primarily determined by the VCO alone is of interest.

III. SPEED-UP TECHNIQUES

Traditional ADC architectures tradeoff precision, bandwidth, sample rate, conversion efficiency, area, and so on. Notable architecture examples include flash ADCs, folding/interpolating ADCs, successive approximation ADCs (SAR ADCs) etc. [13], [15], [23]. Hybrid architectures allow greater freedom for optimization of performance or in overcoming implementation challenges [28].

If we make an equivalence between ADCs and CPUs, many of these ADC architectures can be thought of as analogous to single-core CPU architectures: while, internally, parallel functional blocks exist, these ADCs are conceived to perform single-thread data conversion. But the conversion rate of many of these "single-core" ADCs can be sped up. Two techniques for that are discussed in this section. Their applicability to multiple ADC architectures has been demonstrated.

A. PIPELINING

The first and most mature technique is *pipelining* [16], [29]. To some, *pipelined ADC* is synonymous and exclusive of a particular *architecture* shown in Fig. 6 as an example. While it is, in fact, a much broader technique allowing to speed-up vastly different types of ADC and DAC architectures.

In classic multistage/multistep ADCs [23], D_{out}'s bits are progressively and asynchronously resolved. This is the case

in subranging ADCs or the two-step ADCs: a *coarse* digital estimate (MSBs) is obtained and then, with a cascade operation, a *fine* estimate of the remaining bits (LSBs) is resolved [15]. Coarse and fine bits are separately resolved in folding/interpolating ADCs as well [15].

Pipelining augments the asynchronous cascaded operation by introducing queueing to a multistage converter. Without pipelining, the ADC stages determining the coarse and fine bits are merely cascaded to convert a *single* sample until the process is complete. In this case, the time required to complete the conversion T_{conv} determines the sampling period T_s: $T_s = T_{\text{conv}}$. Instead, by pipelining a multistage ADC, the stages are synchronized to concurrently operate on distinct subsequent samples: while a stage performs the partial conversion of $v_{\text{in}}(k)$, with k being the sample's number, the previous stage operates on $v_{\text{in}}(k+1)$, while the subsequent stage processes $v_{\text{in}}(k-1)$.

A traditional four-stage pipelined 8-b ADC is shown in Fig. 6. The first stage MDAC3[3] resolves the two most significant bits $D_3 = [d_7\ d_6]$, the second stage (MDAC2) determines the next two bits in descending weight's order, $D_2 = [d_5\ d_4]$, followed by the third one (MDAC1) returning $D_1 = [d_3\ d_2]$, and ending with the last stage (Flash) returning the LSBs $D_0 = [d_1\ d_0]$. The ADC's output is $D_{\text{out}} = [d_7\ d_6\ d_5\ d_4\ d_3\ d_2\ d_1\ d_0]$. When ϕ_{even} is asserted, the input v_{in} is sampled on the front-end SHA's output voltage v_s. On ϕ_{odd}, v_s is coarsely quantized into D_3 by the *sub-ADC* below the SHA's output, converted back to analog by the cascaded *sub-DAC* and subtracted from v_s, developing the analog quantity v_{res3}, called the (first stage's) *residue*. Since D_3 is a coarse representation of v_{in}, the residue is the analog fraction of the stage's input that still needs to be quantized. This residue is quantized by the subsequent stage MDAC2. v_{res3} is amplified, in this case with a gain equal to four,[4] to fit the next stage's input range that acquires it through its SHAs when $\phi_{\text{odd}} = 1$. Then the exact same processing is performed by stage MDAC2, resolving the next two significant bits D_2, and producing a *finer* residue v_{res2} when ϕ_{even} is asserted. Note that, while this is happening ($\phi_{\text{even}} = 1$), MDAC3 is acquiring the next v_{in} sample, to be quantized (D_3) as ϕ_{odd} is asserted again. The subsequent stages perform the same operation and, eventually, all bits are determined. The bits determined by each stage are arranged based on their respective weight and combined to return the complete 8b digital output D_{out} as shown in the lower part of the block diagram.

The key point is that, by introducing interstage storage (the SHAs in this example) and synchronous operation, new input samples are acquired and their multistep conversion process can begin *before* the conversion of the previous ones is completed [16], [20], [31]: $T_{\text{conv}} >> T_s$. The completed output D_{out} is returned after LT clock cycles (LT is called *latency*) from when the input $v_{\text{in}}(k)$ has originally been sampled ($LT = 4.5$ clock cycles in this example). By pipelining, the lowest bound on T_s is determined by the slowest stage, not by the overall propagation time through the conversion chain as in an asynchronous multistep ADC.

Numerous variations of this scheme are possible leading to a vast range of architectural variants, hybrids, and design options and tradeoffs: these can differ on the number of stages, the number of bits that each stage resolves (often different from stage to stage and/or augmented by redundancy to tackle nonidealities), the gains used to amplify the residues and the architecture of the sub-ADC.[5]

As hinted, the classic architecture shown in Fig. 6 is *not* the only implementation of pipelining [16], [29]. Pipelining is a *broadly applicable* technique that has been successfully applied to speed-up many types of multistage ADCs, such as: two-step and subranging [32], [33], folding-interpolating [34], various types of discrete-time $\Delta - \Sigma$ ADCs [35], loop-unrolled SARs [36], continuous-time [37] and time-domain ADCs [38]. At the outset, none of these pipelined ADCs look like the architecture of Fig. 6. In fact, pipelining has been used to speed up digital-to-analog converters as well [39], [40]. Abstracting pipelining as a general algorithmic ADC speed-up technique provides additional degrees of freedom in ADC architecture design as also further discussed in Section IV.

B. OPEN LOOP AMPLIFICATION WITH DIGITAL ASSISTANCE

Closed-loop analog circuits are traditionally employed to design ADCs owing to many well-known benefits, including high linearity, insensitivity to PVT drifts, sometimes low noise, etc.

However, when very wide-band performance is required and process technology parasitics are non-negligible, achieving desirable closed-loop stability, power efficiency, or both can be quite challenging. In some cases, the same nominal analog functions can be implemented using faster, or lower power, *open-loop circuits* at the cost of degraded linearity, increased parametric sensitivity, inaccurate gain, etc. But these analog circuits may be complemented by digital circuits that correct for the effects of the open-loop shortcomings as an alternative to feedback. The latter is only a particular instance of a vast class of techniques known as digitally assisted analog (DAA)

3. The acronym MDAC stands for Multiplying Digital-to-Analog Converter [15], [16], [23], [29]. It formally refers to a class of switch capacitor circuits implementing the combination of the pipeline stage's sub-DAC, the residue synthesis, amplification and hold function: all but the quantizer and the local data handling. Commonly, the terms MDAC and pipeline stage are liberally used as synonyms, implying the distinction based on context. This should prepare the reader interested in deeper literature study.

4. Since two bits have been resolved, the residue has a range that is $2^2 = 4$ times smaller than the first stage's input range.

5. While a flash A/D is often used for simplicity and high speed when few bits need to be resolved, it is not uncommon to use a SAR A/D when more bits are quantized and corresponding speed, area or power efficiency meet the requirements.

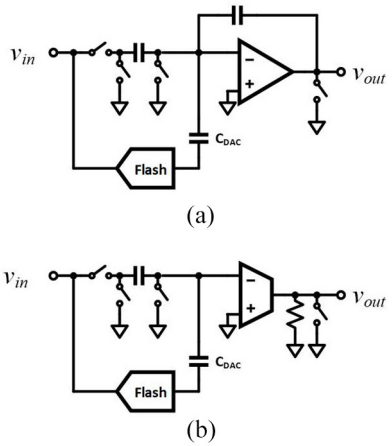

FIGURE 7. (a) Closed-loop MDAC. (b) Open-loop MDAC.

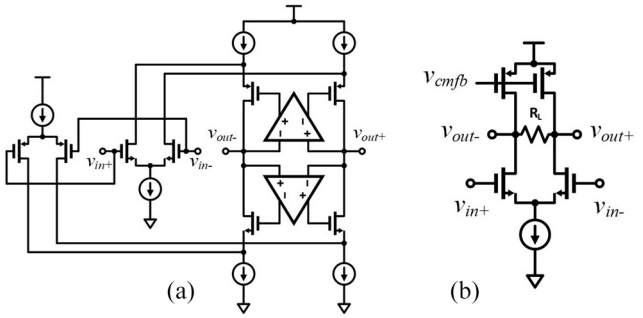

FIGURE 8. (a) High-gain multistage amplifier. (b) Low-gain wide-band amplifier.

design [20], [36], [41], [42], [43], [44], [45], which is process scaling-friendly and that is developing rapidly.

A simplified high-level diagram illustrating a well-known MDAC stage (e.g., like those used in the scheme of Fig. 6) with closed-loop residue generation and amplification is depicted in Fig. 7(a). The residue is obtained by subtracting the analog quantized input from the sampled input itself. The operation is performed by the sampler (shown in the upper input signal path) holding the input and the capacitive DAC controlled by the flash ADC that quantized the input. The difference, at the summing node, is magnified via charge transfer in the feedback capacitor by the closed-loop amplifier and is supplied as an input to the next stage.

This approach has all the known benefits of feedback circuits with respect to accuracy, linearity, noise, etc. But its ability to complete the residue synthesis with enough accuracy and precision in the available time is limited by the gain-bandwidth product of the amplifier, the feedback factor of the switched cap circuit, and by the relative position of all nondominant (parasitic) poles and zeros. The more bits we try to resolve with this stage, the lower the feedback factor (which also degrades because of multiple parasitic capacitors that add up at the summing node).

One way to possibly accomplish the same function *faster* is to use an open-loop residue amplification as the one shown in Fig. 7(b). Here, the *gm-R* circuit performs the amplification, possibly faster, particularly when multiple bits need to be resolved at once. This is possible due to lower capacitive loading at the summing node and the simplified active stage's architecture as shown below. But doing away with feedback means that multiple nonidealities leading to inaccurate gain and to nonlinearity need to be dealt with.

More architecture streamlining, trading-off analog precision and accuracy for speed, is applied to the active circuitry of the MDACs. A high-gain amp intended for closed-loop operation is shown in Fig. 8(a). This is limited in voltage headroom and has multiple nondominant poles/zeros affecting its closed-loop behavior. The simpler *gm-R* stage in Fig. 8(b) is compact, unburdened of most poles/zeros associated with internal nodes, has better voltage headroom, but suffers from higher nonlinear distortion and other nonidealities that need to be compensated for, via specialized digital calibration and linearization, i.e., by DAA [20], [43]. By identifying the source of nonlinear distortion (and its parameters) within this circuit block, it is conceivable to digitally invert the nonlinearity, amend the digitized output and, thus, linearize the conversion [44], [45].

This process is known as *digital post-distortion* linearization [42]. Examples of the use of digitally assisted open-loop amplification stages can be found in widely different ADC architectures such as those in [20], [43], and [93].

It should be noted that the design of such minimalistic digitally-assisted analog circuits requires tremendous analog design mastery as the effectiveness of the digital assistance is intimately tied to the understanding and the ability to properly model the analog behavior that needs correction.

In fact, instead of devising digital assistance as a mere compensation for the resulting analog impairments, an emerging approach is to deliberately *conceive* analog architectures to be well *suited for* digital performance enhancement. The implicit assumption, besides the ability for the digital calibration to be able to perform as needed, is that the added power and complexity of the digital calibration circuits are not voiding the benefits gained by modifying the analog part. Once again, this is where a combined analog-algorithm savviness is key since every case is different and many details matter. As mentioned, technology adoption of high-speed ADCs and integration into ever finer lithography are inseparable. The broad applicability and the scalability of digitally assisted open-loop amplifiers is an important enabler.

As much as one can simplify analog and digitally calibrate, ultimately the process technology is going to limit performance, no matter how much power is spent. This takes us to the next topic.

IV. TIME INTERLEAVING

A seminal paper by Black and Hodges [46] proposed to time-multiplex an array of N identical parallel sub-ADCs to

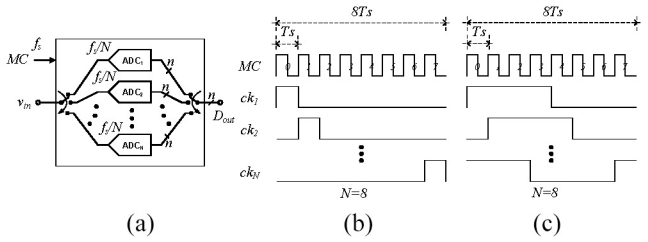

FIGURE 9. (a) TI ADC. (b) Nonoverlapping sampling; $C_{in} = C_s$. (c) Overlapping sampling; $C_{in} = (N/2)C_s$.

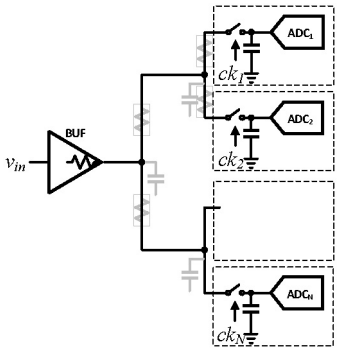

FIGURE 10. Direct interleaving.

effectively sample input v_{in} at a rate that is N times higher than each individual sub-ADC's sample rate. As illustrated in Fig. 9, ADC_1 acquires the first sample, then ADC_2 acquires the second sample T_s seconds later, then ADC_3 acquires the third sample T_s seconds after that and so on. After conversion time, each sub-ADC returns its digital output, so that every T_s seconds a new converted output is available. The sub-ADCs' outputs are time-demultiplexed, assembling the output data stream D_{out} in the original input sampling order. The combination of this array of identical sub-ADCs behaves equivalently to a single ADC sampling N times faster than each of its sub-ADCs.

Fig. 9(b) and (c) show two possible sampling sequences for $N = 8$ channels [47], [48], where MC is the master clock, with frequency f_s, while ck_1-ck_N are the sampling phases associated with the N sub-ADCs at frequency f_s/N. In Fig. 9(b), only one sub-ADC samples the input at any time with a capacitance C_s. In Fig. 9(c), half of the sampling caps C_s are connected to the input at any time, for an individual sampling time of $4T_s$ seconds. In the following, we will always refer to the nonoverlapping scheme of Fig. 9(b).

The powerful and simple idea of TI might lead some to believe that a straightforward bottom-up implementation, starting from designing a robust sub-ADC and replicating it N times, would seamlessly open the way to ultrafast analog-to-digital conversion. Yet, fully integrated implementations have only become feasible in the last ten years or so due to several challenges. One of these challenges is that, once the digital outputs get time-demultiplexed back into a single data stream, the mixing between the effect of the inevitable mismatches among the sub-ADCs (and other interleaved building blocks), together with the processed input sequence, introduces non-negligible spurious spectral content on the resulting demultiplexed output's spectrum. Managing these nonidealities and compensating for their effects has only recently become practical enough to lead to mainstream technology adoption in the form of large-volume commercial integrated self-calibrated time-interleaved ADCs (TI ADCs).

A. ADVANTAGES OF TIME-INTERLEAVING

Other than the benefit of sample rate increase there are other advantages brought by TI.

First, the TI ADC's noise benefits from implicit processing gain. Since one sub-ADC samples the input at any given time, the SNR of the TI ADC, $\text{SNR}_{\text{dB}}(D_{\text{out}})$, is equal to the SNR of the individual sub-ADC $\text{SNR}_{\text{dB}}(D_{o,k})$. Hence, as the TI ADC's sample frequency f_s is N times larger than the individual sub-ADC's sample rate f_s/N, the TI ADC's NSD improves N times

$$\text{NSD}_{dB}(D_{\text{out}}) = \text{NSD}_{dB}\big(D_{o,k}\big) - 10\log_{10}N. \tag{1}$$

Stated differently, there is a processing gain equivalent to oversampling by N.

Second, the conversion efficiency benefits as well. A well-known figure of merit for ADC efficiency is the so-called *Schreier's Figure of Merit*, defined as [13], [15], [16], [35]

$$\text{FOM} = \text{SNR}_{\text{dB}} + 10\log_{10}\frac{\text{BW}}{P}. \tag{2}$$

For the TI ADC, the entire first Nyquist range, $BW = f_s/2$, while BW is equal to $f_s/2N$ for the sub-ADCs. Since $\text{SNR}_{\text{dB}}(D_{\text{out}}) = \text{SNR}_{\text{dB}}(D_{o,k})$ the TI ADC's FOM is

$$\text{FOM}_{TI\ ADC} = \text{SNR}_{\text{dB}}(D_{\text{out}}) + 10\log_{10}\frac{f_s/2}{P_{TI}} \tag{3}$$

where P_{TI} is the TI ADC's power consumption. If the power spent in interleaving the sub-ADCs is negligible compared to the total contribution of the power $P_{sub-ADC}$ consumed by each sub-ADC then $P_{\text{TI}} = N{\cdot}P_{\text{sub-ADC}}$, and the TI ADC has the same power efficiency as each one of the sub-ADCs [20]

$$\text{FOM}_{\text{TI ADC}} = \text{SNR}_{\text{dB}}\big(D_{o,k}\big) + 10\log_{10}\frac{N\cdot f_s/2N}{N\cdot P_{sub-ADC}} =$$
$$= \text{SNR}_{\text{dB}}\big(D_{o,k}\big) + 10\log_{10}\frac{f_s/2N}{P_{\text{sub-ADC}}} = \text{FOM}_{\text{sub-ADC}}. \tag{4}$$

This is an important result since a very power efficient sub-ADC can be selected for a low sample rate.

There is an additional aspect to note. Regardless of the sub-ADC's architecture, at low frequency, its power can be generally modeled to be proportional to the sample rate: $P_{\text{sub-ADC}} \approx P_q + \alpha \cdot (f_s/N)$, where P_q is the quiescent power and α is a proportionality constant used to quantify the dynamic power contribution. P_q is generally negligible compared to this latter dynamic power term. Therefore, the logarithmic term in the FOM formula (4) is approximately constant. This proportionality between dynamic power and

sample rate holds until the sample rate approaches process, circuit, and architecture limits. An example of these limits is, if, to expedite settling, the MDAC's amplifier like those in Figs. 7 and 8, is biased for a nonoptimal g_m/I_d. Or, if the MOS channels of some critical devices reach carrier velocity saturation, etc. When operating the sub-ADCs beyond these limits, $P_{sub-ADC}$ is not proportional to the sample rate anymore and it grows much more rapidly with the clock rate. As a result, the logarithmic term in (4) decreases for increasing sample rate, highlighting a condition often referred to as *approaching the technology front* of the FOM.

The sample rate at which the conversion efficiency drops in the sub-ADC is $1/N$ lower than the corresponding TI ADC's sample rate f_s. In other words, the TI ADC surpasses the limit set by the technology front by performing as in (4) but for a sample rate that is N times higher than the sub-ADC.

Unfortunately, as the original simplifying assumptions are carefully considered, limitations emerge. The power spent in the interleaving overhead circuitry is not negligible and grows with the size of the array N. Moreover, some overhead circuits do not operate at a fraction of f_s as the sub-ADCs. For instance, front-end input multiplexing circuitry, output demultiplexing circuitry, and some of the clocking circuitry are clocked either at the full rate f_s or at a larger fraction than f_s/N. Therefore, for large f_s, their dynamic power grows with f_s^{β} with $\beta \geq 2$ [49], [50].

As a result, (4) needs to be modified as

$$\mathrm{FOM_{TI\ ADC}} = \mathrm{FOM_{sub-ADC}} - \mathrm{OP}(N, f_s) \tag{5}$$

where *OP(N, f_s)* is an amount, closely dependent on the implementation, accounting for the overhead power penalty. This limits the efficiency of interleaving for larger arrays and a growing sample rate.

B. INTERLEAVING ARCHITECTURES

The simplest way to time-interleave N sub-ADCs is by direct multiplexing, as shown in Fig. 10, where the input signal, possibly after buffering, is directly supplied to the sub-ADC array. Examples of *direct interleaving* are common [21], [22], [34], [38], [50], [51], [52], [53] for a relatively low (e.g., $N = 8$) number of sub-ADCs. As the diagram in Fig. 9 suggests, a large N, among others, results in a large signal distribution tree with non-negligible performance-limiting parasitics.

Design compromises are possible by recursively time-multiplexing with a *hierarchical interleaving* architecture as the one depicted in Fig. 11. A first interleaving rank of L samplers/SHA ($L = 2$ in Fig. 11) captures the buffered input v_{in} at an individual sample rate f_s/L. In turn, each sampler SHA drives a second rank of K sub-ADCs ($K = 4$ in Fig. 11), at an individual sample rate $f_s/(L \cdot K) = f_s/N$. This approach allows for mitigating the previously highlighted input interconnect limitation and relaxing multiple other frequency-dependent challenges discussed later. It reduces front-end complexity, allowing a higher net interleaving count $N = L \cdot K$.

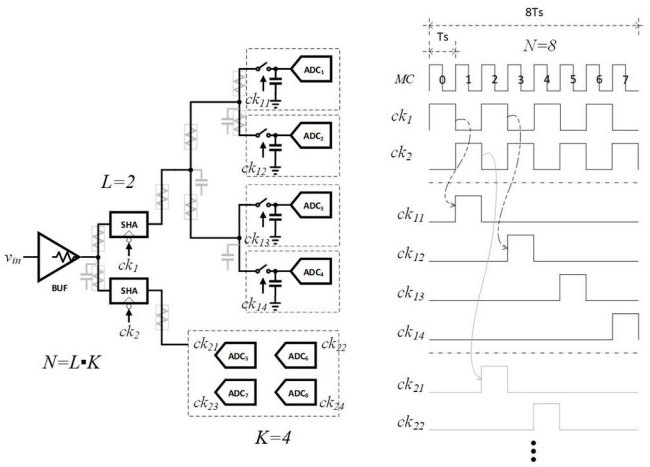

FIGURE 11. Hierarchical interleaving.

L internal signal trees are locally introduced at the output of the first rank. Each tree can be limited in size by the number K of local sub-ADCs and it operates at a fraction f_s/L of the total sample rate.

This reduction in local complexity and additional degrees of design freedom require however the introduction of the intermediate overhead SHA stages. Each one of these stages contributes to more power consumption, noise, and nonlinearity. Highly interleaved ADCs (e.g., $N > 8$) are almost invariably hierarchically interleaved [54], [55], [56], [57], [58], [59], particularly if implemented in fine lithography processes, where large interconnect structures severely limit high-frequency performance.

The combination of pipelining and TI offers another degree of freedom in ADC architecture design and optimization. Because while the design of the first stages of the sub-ADC is far from trivial, pipelining a core sub-ADC allows using a lower number N of sub-ADCs. Multiple examples of that are provided in Section VI and Table 4.

C. INTERLEAVING ARTIFACTS

The biggest challenge with interleaving is that mismatches among the interleaved samplers and quantizers constituting the TI ADC introduce spurious signals in the demultiplexed output D_{out}.

Assuming a sinusoidal input v_{in} at frequency f_{in}, and that all sub-ADCs have ideal samplers and quantizers and, all N of them, referring to Fig. 9, are identical to each other, the output spectrum of D_{out} consists of a single tone (v_{in}), plus noise and quantization, as in Fig. 12(a).

However, mismatches in: 1) offset; 2) gain error; 3) sampling time skew; and 4) bandwidth among the time-interleaved samplers and quantizers exist. As a result, the corresponding output spectrum of D_{out} shows additional images of the input tone at different frequencies and power

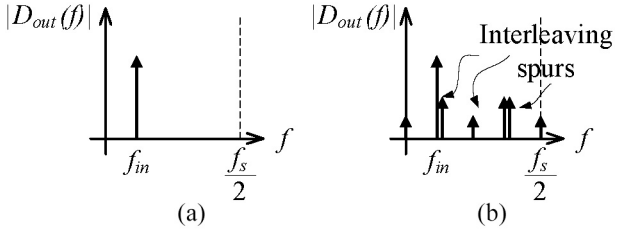

FIGURE 12. (a) Ideal TI ADC output spectrum. (b) Actual TI ADC output spectrum.

levels as depicted in Fig. 12(b). These spurs are known as *interleaving spurious signals* or *TI spurs* or *artifacts*.

Each of the four mentioned impairments impacts all converted codes. However, if the impairments were identical for all the output codes D_{out} their net effect would be the same as in a traditional noninterleaved ADC, merely introducing a dc offset and a gain error [23].

If there is a mismatch among the sub-ADCs, the codes in the demultiplexed output series D_{out} experience periodic errors. For instance, if code $D_{\text{out}}(k)$ is converted by sub-ADC j (with $j = 1, \ldots, N-1$), then codes $D_{\text{out}}(k+N)$, $D_{\text{out}}(k+2N)$, $D_{\text{out}}(k+3N)$ and so on are affected by the sub-ADC j impairments. The next code $D_{\text{out}}(k+1)$, converted by sub-ADC $j+1$, as well as codes $D_{\text{out}}(k+1+N)$, $D_{\text{out}}(k+1+2N)$, $D_{\text{out}}(k+1+3N)$ are affected by the sub-ADC $j+1$ impairments, which are mismatched (different) from those associated with sub-ADC j. The process can be repeated for all converted data and this effect can be thought of as discrete-time mixing between the N impairment sets of the sub-ADCs and the sampled data series.

The effects of these TI impairments have been extensively modeled, formalized, and quantified both deterministically and stochastically [60], [61], [62], [63], [64], [65]. An intuitive introduction of each one is provided here, deferring the reader to the literature for in-depth coverage.

Offset mismatches result into an N periodic sequence of N different offsets superimposed to D_{out}, adding fixed tones at[6]:

$$f_{o,k} = k \cdot f_s/N, k = 0, 1, 2, \ldots \tag{6}$$

accounting for the repetition period N of the mismatch error pattern and its harmonics and aliases. The powers of these spurious tones (excluding the dc term) depend only on the mutual mismatches between the offsets. The power and fixed-frequency location of these tones is completely independent of the input signal. In fact, these tones are present even in the absence of input v_{in}. This is a static impairment.

Gain error mismatch is equivalent to amplitude modulation of the converted series D_{out} by the N-periodic gain error sequence, adding power-scaled images of the original input to the D_{out} spectrum at[7]

$$f_{g,k} = \pm f_{\text{in}} + k \cdot f_s/N, k = 1, 2, \ldots \tag{7}$$

6. $k = 0$ corresponds to dc, namely, to the TI ADC's offset.

7. In this case k starts at 1, since (7) gives the frequency of the interleaving images. Setting $k = 0$ would give the frequency of the input signal itself.

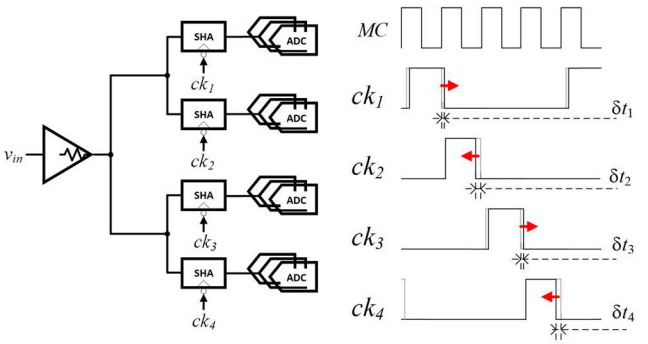

FIGURE 13. Sampling time skew.

These spurs are dependent, in power, on both the converted signal's power and the periodic gain error mismatch series' power. The power of the images does not depend on f_{in}. Because of that, this is also a static impairment.

Sampling time skew occurs at the front-end sampling stage, as shown in the example for $L = 4$ samplers of Fig. 13. The actual sample times, determined by the falling edges of ck_1, ck_2, ck_3, and ck_4 are displaced with respect to the uniform sampling master clock MC by time skews δt_1, δt_2, δt_3, and δt_4, respectively.

As with gain error mismatch, this periodic error pattern mixes with the input sampled data series. In this case, however, the effect is reminiscent of sampling aperture jitter described in Section II-B, in the sense that each time skew δt_{1-4} adds a voltage error δv_{1-4} to the corresponding sampled input v_{s1-4}. Unlike jitter, in the case of sampling skew, the error sequence due to δt_1, δt_2, δt_3, and δt_4 is L-periodic and, hence, its effect is equivalent to phase modulation of the converted series D_{out} by the L-periodic timing error sequence, adding power-scaled images of the original input to the D_{out} spectrum at

$$f_{t,k} = \pm f_{\text{in}} + k \cdot f_s/L, k = 1, 2, \ldots \tag{8}$$

The power of timing skew artifacts depends on both the converted series, the input frequency f_{in}, and the periodic skew series. Note that the set of image frequencies in (8) is a subset of (7) for a hierarchical interleaving architecture ($L < N$) while it coincides with (7) for direct interleaving ($L = N$). Also note that, for analogous reasons as sampling aperture error, the spurious power introduced by this impairment, decreases[8] when f_{in} decreases and, hence, it is a dynamic impairment.

The last type of impairment causing spurious input images is bandwidth mismatch. Like sample time skew, bandwidth mismatch also occurs at the TI ADC's input as shown in Fig. 14. Fig. 14 shows an example of hierarchical interleaving with $L = 2$ that suffers from bandwidth mismatch but no sampling time skew.

8. In the extreme case in which f_{in} approaches dc, a time-displacement δt_{1-4} over nearly constant v_{in}, causes no appreciable deviation δv_{1-4} and hence the power of the corresponding spurious images vanishes (unlike the gain error images).

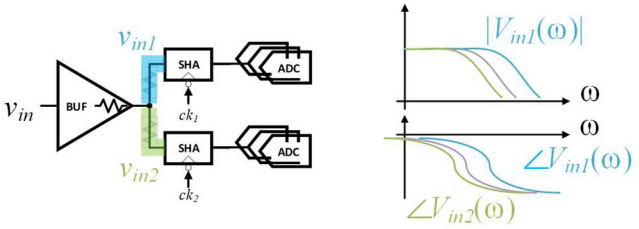

FIGURE 14. Bandwidth mismatch.

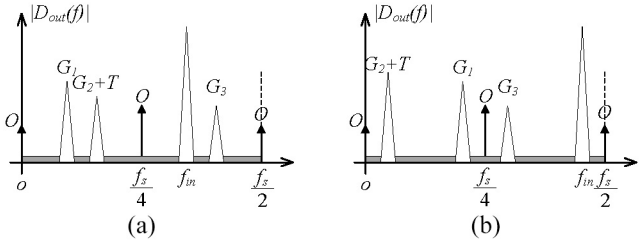

FIGURE 15. Examples of spectra for $L = 2$ and $N = 4$. (a) Mid-range f_{in} case. (b) High frequency f_{in} case.

TABLE 2. TI artifacts.

Mismatch	Spurious Frequency	Spurious Power Dependence
Offset	Depends on N. Independent from input. Eq. (6)	Independent from input
Gain error	Depends on f_{in} and N. Eq. (7)	Grows with $P(v_{in})$. Independent from f_{in}.
Timing skew	Depends on L and f_{in}. Eq. (8)	Grows with $P(v_{in})$ and f_{in}.
Bandwidth	Depends on L and f_{in}. Eq. (9)	Grows with $P(v_{in})$ and f_{in}.

The input signal $v_{\rm in}$ is buffered and fed to the two identical samplers through a signal tree and should result into two nominally identical SHAs' input voltages v_{in1} and $v_{\rm in2}$, respectively. The combination of the buffer's output impedance, the interconnects parasitics between buffer and SHAs, SHAs' sampling network strays, and sampling capacitors determine the frequency response between the common buffer output and v_{in1} and $v_{\rm in2}$, respectively.

These two frequency responses should be identical and with the same bandwidth $\omega_{\rm 3dB} = \omega_{1,\rm 3dB} = \omega_{2,\rm 3dB}$. Though due to mismatches in strays, on the interconnects, or the on-resistance of the two sampling networks or the two sampling capacitors, frequency-dependent mismatch between $V_{in1}(\omega)$ and $V_{in2}(\omega)$ arise and result in $\omega_{\rm 3dB} \neq \omega_{1,\rm 3dB} \neq \omega_{2,\rm 3dB}$. That is equivalent to say that v_{in1} and $v_{\rm in2}$ are affected by two different gain errors and two phase shifts upon sampling, even in the absence of previously considered TI ADCs errors. Moreover, these errors on v_{in1} and $v_{\rm in2}$ are frequency dependent and input images are introduced, just like in the case of gain error mismatch and timing skew mismatch, at

$$f_{bw,k} = \pm f_{\rm in} + k \cdot f_s/L, k = 1, 2, \ldots \tag{9}$$

In many practical cases, $\omega_{\rm 3dB} >> 2\pi f_{\rm in}$. Therefore, any equivalent gain mismatch between the two paths is largely negligible in the $f_{\rm in}$ range. But notable phase shifts emerge rapidly for increasing $f_{\rm in}$, well before approaching $\omega_{\rm 3dB}/2\pi$. Even for moderate $f_{\rm in}$, the phase mismatch between v_{in1} and $v_{\rm in2}$ can be sufficient to lead to noticeable timing errors between the two paths even in the absence of sample time skew (it is the input signal that is skewed differently, not the time of sampling). If the skews were frequency independent, then the net errors could be lumped with sample time skew. But the signal skew due to bandwidth mismatch grows with the input frequency $f_{\rm in}$, causing additional challenges as discussed in the following sections.

A summary of the TI artifacts is reported in Table 2.

Fig. 15 shows two examples of spectra summarizing what is covered above. A narrowband input signal $v_{\rm in}$ with center frequency $f_{\rm in}$ is digitized by a hierarchical TI architecture with a first rank constituted of two SHAs ($L = 2$), each of which drives a second rank of two sub-ADCs ($K = 2$). That is, $N - L \cdot K = 4$ interleaved ADC.

Offset mismatch introduces input-independent tones at frequencies $f_{o,k}$ as from (6), visible at dc, at the Nyquist frequency (f_s/2) and in the middle of the 1st Nyquist band (f_s /4) and marked with an O in Fig. 15.

Gain mismatch introduces three input signal's images, G_1, G_2 and G_3 with center frequencies $f_{g,k}$ given by (7). From (7), it is seen that the frequencies $f_{g,k}$ of the images track $f_{\rm in}$. These images are "reflected" around four subranges (each one with bandwidth f_s/8) of the TI ADC's Nyquist band f_s/2. That is consistent with the fact that each of the four sub-ADC clocks at a quarter of the rate of the master clock. The powers of G_1 and G_3 do not depend on $f_{\rm in}$ (although their frequency does). Their power only depends on the input signal power and the respective gain error mismatches.

Timing skew and bandwidth mismatch introduce a single image T (since $L = 2$), whose frequency $f_{t,k} = f_{b,k}$ is given by (8) [or (9)], which coincides with G_2's frequency. In fact, G_2 and T add to create a single image marked as $G_2 + T$, as shown in Fig. 15. Its total power grows with $f_{\rm in}$ since the contribution from the timing skew (and bandwidth mismatch) depends on $f_{\rm in}$ and it adds to the frequency-independent contribution of the gain error mismatch. This is visible comparing Fig. 15(a) and (b): as $f_{\rm in}$ increases, the frequency of image $G_2 + T$ decreases (reflected around f_s/4 because of the front-end 2x SHA multiplexing) but it grows in power because T's power grows as a result of the increase in $f_{\rm in}$.

V. INTERLEAVING CALIBRATION AND RANDOMIZATION

Since the above mismatches cause undesired images on the ADC output spectrum, the question is what can be done to eliminate them or to minimize them? Besides obvious careful circuit and layout design, the answer is calibration.

There are many ways to calibrate for the mismatches discussed in the previous section. Each approach comes with compromises. Before diving into different calibration

algorithms and circuits, it is appropriate to refresh related taxonomy and nomenclature.

We start by postulating a suitable mathematical model for the ADC and its nonidealities. A parametric model grounded on a clear understanding of this physical system is highly appropriate.

The act of determining the numerical values of the model's parameters is called *system identification*, or *estimation*. To identify the model's parameters, stimuli are applied, and the system's response is sensed. If the stimuli are especially applied to estimate, the term *training* stimuli is used; some algorithms do not require them and estimate based only on the regular inputs. If the model's parameters, sometimes called "the coefficients," can be identified, then this information can be used to either take a *corrective* action *over the nonideality* (on the cause) or to cancel *its effects* on the ADC's output.

The estimation can be done in the analog domain. This often requires introducing analog circuitry to sense *the nonidealities*, right where they happen.

However, sensing circuitry's accuracy or precision can be limited to keep it simple. The latter is often a critical requirement because, besides adding overhead, sensing circuitry inevitably perturbs the sensed circuit, possibly making the nonideality or other performance aspects worse.

Estimation can also be performed digitally. Often this requires assessing *the effects* of the nonidealities, after the conversion (or part of it).

Estimating the causes based on their effects can be hard. It may be subject to latency or it could require high clock rate logic, possibly adding considerable power consumption.

Assuming that nonidealities can be estimated using analog or digital circuits, the estimated values are used to compensate for the *effect* of these nonidealities so that the resulting digitized signal more closely resembles its analog value. This process is known as *correction*. Correction can be implemented as analog or digital correction, with similar tradeoffs as the one mentioned above for estimation. Analog correction is used to directly counter nonidealities, but the *insertion* of correction circuits can introduce problems as well.

Calibration algorithms are categorized as *foreground* or *background* calibration. Foreground calibration is performed *before* the ADC is ready to be used, for instance, at power up or while the ADC is idle. Background calibration is performed while the ADC is converting a signal, during regular operation. If the estimated parameters/nonidealities vary during normal operation, background calibration senses the changes and adapts to them. Examples of such variations are those caused by temperature changes or parametric drifts due to device aging. For some background calibration algorithms, the regular input signal exercises the nonidealities requiring estimation. Alternatively, additional *training* stimuli are injected in the presence of the converted signal to exercise the nonidealities together or independently from the converted input. Often, the additional stimulus is pseudo-random noise, that is subsequently digitally removed from the calibrated ADC output.

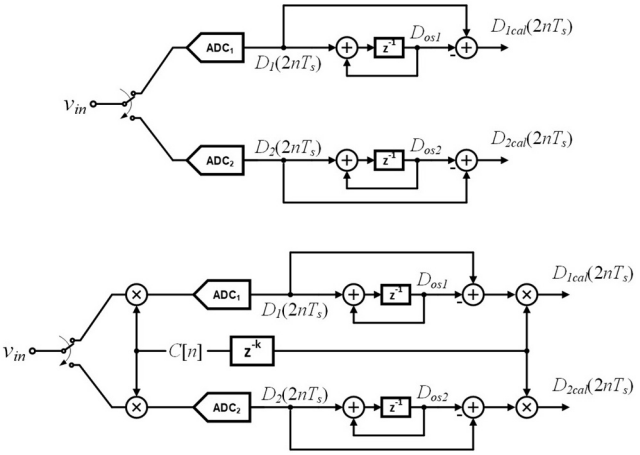

FIGURE 16. Examples of digital offset mismatch calibration [16], [66], [68].

In the remainder of this section, representative examples of TI calibration algorithms are presented. In general, robust solutions for offset mismatch calibration and gain error calibration are available. Timing skew calibration is more challenging and while multiple algorithms exist, each has specific advantages and drawbacks. Finally, bandwidth mismatch calibration is very challenging. Known solutions that tend to work well under specific assumptions on the input signal or the application, often also add large area or power overhead. It could not be stressed enough that, given the importance of this topic, intense research is ongoing and new and improved algorithms are emerging at an impressive pace, relaxing some of the existing tradeoffs.

A. OFFSET MISMATCH CALIBRATION

An example of an offset mismatch calibration scheme is shown in the upper half of Fig. 16 [16], [66]. A 2-times interleaved ADC ($N = 2$) is considered for simplicity of explanation, but this can be generalized to more channels.

The input signal v_{in} is used as an estimation stimulus and this digital calibration runs in the background.

In Fig. 16, v_{in} is digitized by the ADCs and the integrators that low-pass filter the outputs[9] of each sub-ADC, determine the respective dc components D_{os1} and D_{os2} that estimate the offsets. These are subtracted from the sub-ADC outputs, to obtain the calibrated outputs $D_{1\,cal}$ and D_{2cal}.

This scheme introduces frequency nulls to the demultiplexed output spectrum where the offset spurs would otherwise emerge. The dc inputs of each sub-ADC are zeroed and, the sub-ADCs' dc signals coincide with v_{in}'s frequency components at $f_{o,k}$ in (7). Alternatively, this scheme could be modified to force the outputs to have equal offsets.

9. Digital low pass filters are used in actuality. The diagrams in Figs. 16 and 17 are simplified to ease explanation.

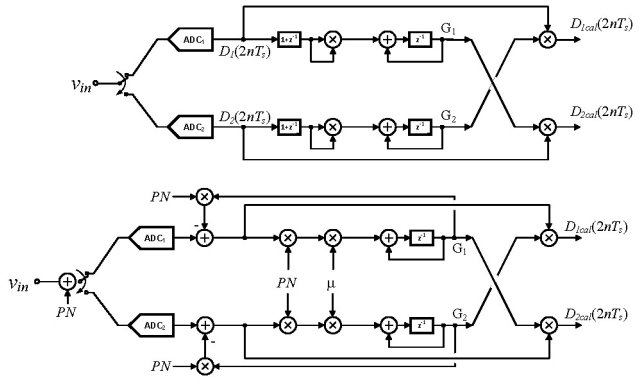

FIGURE 17. Examples of digital gain error mismatch calibration [16], [66], [68].

The input signal v_{in} could be separated from the dc offsets of the sub-ADCs by modifying the previous scheme as shown in the bottom half of Fig. 16. A chopping sequence $C[n]$ is introduced, randomly alternating between $+1$ and -1 (a PN sequence), so the sub-ADCs do not see constant series when the input v_{in} is at the notch frequency $f_{o,k}$.

B. GAIN ERROR MISMATCH CALIBRATION

The scheme on the upper half of Fig. 17 estimates the gains of each of the two sub-ADC by computing their average output power. Ideally, these two averages should be equal, so any differences are an indication of the gain mismatches.

Let us consider the upper sub-ADC path and ignore the $1 + z^{-1}$ filter first. The sub-ADC's output is squared (see multiplier with equal inputs) and low-pass filtered by the integrator, whose output gives a real time estimate of the signal power G_1. This value is used to scale the output of the other sub-ADC. As a result, the calibrated output of each sub-ADC has the same average gain $G_1 \cdot G_2$. Note that G_1 and G_2 can be normalized to be close to 1.

The filter $1+z^{-1}$ notches the Nyquist frequency of the sub-ADC. Similar to the previous offset calibration scheme, if the input signal v_{in} lies at some specific frequencies discussed later, the algorithm estimation diverges even without gain mismatch. While this filter does not affect the signal, it prevents the estimation from diverging [67], [68].

However, if the input v_{in} is not active for some time or it is small, the accuracy of the estimates for the gains degrades. Once the input is large again, the impact of these inaccuracies emerges, and artifacts resurface. It takes some time for G_1 and G_2 to reconverge to accurate gain estimates.

A recurring theme should have been noticed by now. Namely, algorithms that rely on the input signal as a stimulus and estimate nonidealities solely based on the sub-ADC outputs often suffer from convergence issues due to pathological input signals. This issue is elaborated on later.

To counter that, a modified scheme is reported in the lower half of Fig. 17, where a pseudo-random noise sequence *PN* is added to the input v_{in}. The *PN* stimulus is used to drive the algorithm toward the desired gain estimates, while making it insensitive to v_{in}. A *PN* sequence is used because it is uncorrelated from the input v_{in} and, hence, separable from v_{in}.

The gain of each sub-ADC is estimated iteratively by the least mean square (LMS) algorithm driven by the *PN* stimulus which correlates out the input v_{in}. The LMS algorithm is described by the gain estimation update equations

$$G_i[n+1] = G_i[n] - \mu \cdot PN[n](PN[n]G_i[n] - D_i[n]) \quad (10)$$

where μ determines the update rate of the gain estimate.

Once again, the gain estimate is average output power. Equation (10) correlates with *PN*, driving the gain estimates, but correlates out the sub-ADC's outputs representing v_{in}.

Note, in the lower scheme shown in Fig. 17, that the *PN* is subtracted from the sub-ADC outputs, which are then equalized with the gain estimates as in the previous scheme, and a *PN*-free calibrated output is returned on the right [16].

C. TIMING SKEW CALIBRATION

The first approach, shown in Fig. 18, employs a separate ADC, marked as ADC_R, known as *reference ADC*.

The N-channels TI ADC ($N = 4$ in this example) that needs calibration is shown below ADC_R. ADC_R subsamples v_{in} together with the TI ADC. ADC_R's output is compared with the sub-ADCs output and this information is used to deskew each sub-ADC's sampling instant. Suppose ADC_R and one of the sub-ADCs, let us call it ADC_A, are simultaneously sampling v_{in}. Assume their sampling times are determined by the falling edges of their respective sampling clocks ϕ_{cal} and ϕ_A. If ϕ_{cal} and ϕ_A are time-aligned (no skew), then the respective outputs are equal. Otherwise, the outputs are going to be slightly different, and this information is used to align ϕ_A to ϕ_{cal}.

If ADC_R is clocked at full-rate f_s and has the same resolution n as the interleaved ADC, the output comparisons are trivially implemented to adjust the skews. But we would not need the TI ADC since ADC_R would already be capable of doing its job all by itself. So, the question is what type of algorithm can zero out all the N skews by using a minimal ADC_R and require the least overhead? A metric of the output code difference between the reference and the calibrated ADC is required to digitally control the falling edge of ϕ_A until the metric is optimized.

Let us look at the time diagrams in Fig. 18(b). If ADC_R clocks at $f_s/(N+1)$, then ADC_R samples v_{in} at about the same time as ADC_1. Next, all other N-1 sub-ADCs sample v_{in}, though not ADC_R. Then, ADC_1 samples v_{in} again. Next, ADC_R samples v_{in} at the same time as ADC_2. If the sampling times are aligned, then ADC_R and ADC_2 should have the same output.

ADC_R's sample rate and the sub-ADC's sample rate are relatively prime ($N/(N+1)$) so ADC_R periodically samples together with each of the sub-ADCs at a regular $N(N+1)T_s$ cadence.

A possible metric to be used to zero out the N time skews is the cross-correlation between the reference ADC_R

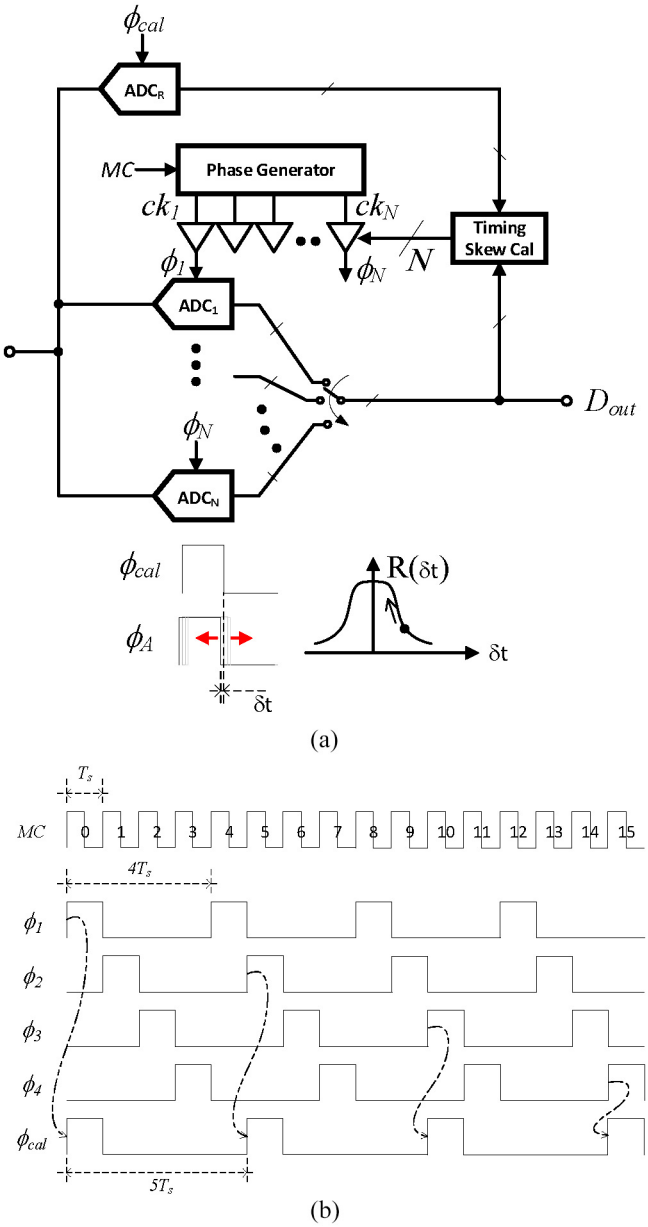

FIGURE 18. Example of mixed-signal timing skew calibration [74], [75]. (a) Block diagram. (b) Clock phases.

and each sub-ADC. That is, the N cross-correlation functions between the output code series realized using the ADC$_R$ outputs and the corresponding sub-ADCs outputs when simultaneously sampling v_{in}. The cross-correlation $R(\delta t)$ used for ADC$_A$, is exemplified in Fig. 18(a) where δt is the estimated timing skew requiring correction. The skew is zeroed out when the cross-correlation is maximized. $R(\delta t)$ drops if ADC$_A$ drifts with respect to ADC$_R$ and their respective outputs and sampling times move away from one another.

The correction algorithm individually digitally controls the N phases ϕ_1-ϕ_N, based on the N cross-correlations, aligning the skews to the reference set by ADC$_R$ and ϕ_{cal}.

In many implementations, the reference ADC$_R$ can be reduced to a single comparator, minimizing the overhead and leveraging the lengthy time series required by the cross-correlation functions.

The challenge common to methods relying on cross-correlation is that for a wide-band input, or a somewhat random input waveform, the correlation function has a shallow slope and so the algorithm convergence time can be very slow and have a strong dependence on the nature of v_{in}. Moreover, the loading to the input signal source is inconsistent. When ADC$_R$ samples v_{in} together with one of the sub-ADCs, two samplers are simultaneously loading the input source. That is the case only once every $N + 1$ samples, not for every sample. This affects the sampled input itself and introduces a calibration error, introducing spurious artifacts on the output spectrum [75].

The principle behind the described algorithm is common to many others, with different advantages and disadvantages. For instance, in [76], ADC$_R$ is replaced by a so-called *window detector* which is a comparator sampling the input at full rate f_s, flagging if v_{in} is (within a narrow range from) crossing a prescribed threshold value. When flagged, the output code of the sub-ADC sampling v_{in} is saved. The saved output codes will statistically accumulate above or below the set threshold depending on whether the sub-ADC's sampling edge is leading or lagging the window detector. This estimation directs the correction for the sampling edges.

This approach solves the problem of the inconsistent input source loading since it samples it at every T_s. Moreover, using the statistics on threshold crossing instead of the cross-correlation, it has been shown to converge faster, particularly for an active input that crosses the threshold often. However, it has its own challenges that are, for instance, associated with the window detector's design and dependencies on the choice of the window's size or the threshold [76]. Moreover, since the window detector samples at f_s, a very high sample rate implementation is challenging.

Another interesting example is described in [77], using two reference ADCs. One is preceded by an analog high pass filter to determine the slope of v_{in} while the second one is the usual timing-skew-free reference. While the v_{in}'s slope digitization allows for a faster and better estimation of the skews, this supplementary ADC adds to the overhead and input loading.

There are other methods that, instead of relying on a separate reference ADC, determine the skews by monitoring the outputs from sample to sample and from sub-ADC to sub-ADC. They refer to one another sub-ADC outputs, instead of a common reference, to determine any *mutual* time errors. This is the general idea behind the algorithms in [21], [48], [71], and [86].

In all examples thus far, the correction part of the calibration has been realized by digitally controlling the position of the sampling edges. But in [78], [79], [80], [81], [82], [83], [84], [85], and [86], instead of correcting these analog errors, the ADC's output

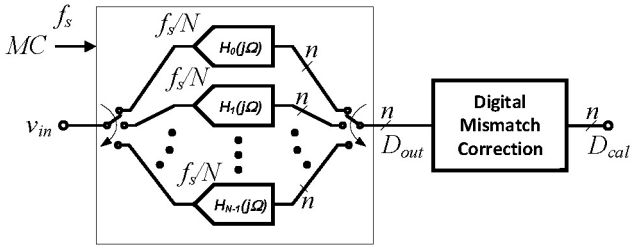

FIGURE 19. Digital mismatch compensation based on a linear time-varying model [78].

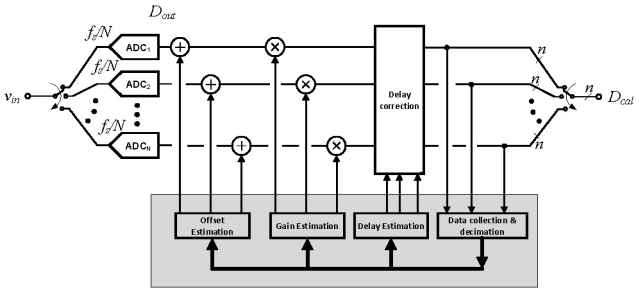

FIGURE 20. Compensation using blind identification and digital correction [70].

codes are digitally compensated to cancel the effects of the skews. More discussion of some of these algorithms follows.

D. BANDWIDTH MISMATCH MITIGATION AND COMPENSATION

Bandwidth mismatch is recently emerging as a serious concern in TI ADCs. The sample rate of TI ADCs is rapidly increasing and so is the input signal v_{in}'s frequency f_{in}, but improvements on parasitics are not keeping up. Therefore, it is not viable to mitigate this by maximizing the input bandwidth ($\omega_{3\text{dB}} >> 2\pi f_{\text{in}}$).

A different approach has been proposed in [20], where the front-end has two interleaved samplers driving a sub-ADC array. The samplers' mismatch is mitigated by introducing a third sampler and by randomly selecting the next available sampler at each sample time. In this way the periodic error introduced by the bandwidth mismatch is randomized and the power of the corresponding artifacts is reduced, trading off SFDR for noise power (increased by the randomized spurious power).

E. DIGITAL COMPENSATION BASED ON A TIME-VARYING MODEL

Compensation for all TI mismatches, including bandwidth mismatch, is directly addressed in [78], [79], [80], [81], [82], [83], [84], and [85]. Referring to Fig. 19, the general idea in these digital compensation approaches is that *each* sampler (and sub-ADC) can be modeled by its continuous-time linear time-invariant transfer function $H_k(j)$ $(k = 0, \ldots, N-1)$ and corresponding mismatched coefficients, followed by a sampler and quantizer,[10] hence capturing bandwidth mismatch into the model.

Due to interleaving, v_{in} is periodically applied to this bank of continuous-time filters (and samplers and quantizers) and, hence, the TI ADC is effectively modeled as a linear (periodic, or polyphase) *time-varying* continuous-time filter, cascaded by suitable samplers and quantizers.

It can then be conceived to digitally compensate for the errors affecting D_{out} by applying a (interleaved) digital mismatch correction consisting of multiplexing a corresponding bank of inverse filters. Each digital inverse filter equalizes one of the analog mismatched filters (sub-ADCs) so the frequency response of the resulting cascade is consistent for all multiplexed codes and a compensated output data stream D_{cal} is returned.

The main strength of this approach is its fully digital implementation: it scales and avoids inserting sensing or compensation circuitry on sensitive analog circuits.

But multiple signal processing and digital implementation challenges have, so far, limited a wide fully integrated use in commercial ADCs. To cite only a few, the known limitations of equalizing a continuous time filter with a digital one over a wide bandwidth approaching the 1st Nyquist band, and the rapid growth in computational complexity (primarily associated with the number of digital multipliers, their required speed and bit width) as a function of the target SFDR and SNDR [78], [79], [80], [81], [82], [83], [84], [85]. Finally, this type of impairment estimation belongs to the class of so-called *blind identification* algorithms, which has limitations described as follows.

F. BLIND IDENTIFICATION

The process of identifying the multiple parameters of a multidimensional model (the ADC model, in our case) only using its outputs and without requiring input training stimuli is called *blind identification*.

Fig. 20 shows a very general scheme of a TI ADC's blind identification and digital cancelation [78]. The outputs of the sub-ADCs are collected and used to estimate, in general, offsets, gain errors, and skews. Possibly also bandwidth mismatch, though that is omitted in Fig. 20 for simplicity and without loss of generality.

The estimated values are used to correct the data series, starting from the offset, followed by gain and sampling delay errors. The estimation and the corrections to be made do not need to be computed at the same clock rate of the entire ADC or of the sub-ADCs. The computations inside the grey box can be performed at a lower clock rate, possibly after decimation. While this slows down the convergence of algorithms, it reduces digital overhead and power and allows executing complex calculations that are onerous at a high clock rate. The key advantage of digital blind identification algorithms is that they avoid introducing sensing or actuation overhead in the sensitive ADC's circuitry. There are several blind identification algorithms

10. This modeling approach is valid for only a single Nyquist band [78].

capable of calibrating for offset mismatch, gain error mismatch, timing skew, and bandwidth mismatch together [41], [42], [78], [79], [80], [81], [82], [83], [84], [85], [86]. The identification process frequently consists of the optimization of a cost function, often using a correlation function. The cost function's minimum corresponds to the unknown mismatches. The solution algorithm often uses some form of gradient descent LMS algorithm.

However, since a blind identification relies on the ADC outputs only, the algorithm cannot distinguish if a change in the outputs' behavior is due to a change in the inputs or a change in the model's parameters (e.g., the mismatches). A common assumption is that the input v_{in} is quasi-stationary (QS). Namely, that it can be represented as a random process that does not change mean and correlation over time [85]. So, if the output stream starts to change these metrics, one must conclude that this is happening because some PVT variation made the mismatches vary. The algorithm detects the change and adjusts the estimate of the mismatches. That can be a problem because while, *in steady state*, the input is indeed QS, the input may suddenly change to transition to a different mean or correlation at which it eventually stays constant for a while. As described in Section II-A, in a wireless communication system, a channel may turn on or turn off suddenly. So, the ADC's inputs behave as QS until a new channel lineup emerge, transitioning to a different "steady-state" mean or correlation.

This output behavior transition triggers a drift in the mismatches' *estimation*, which *did not* change and should not be reidentified with new values. The corresponding transitory corrective action causes an undercorrection or an overcorrection and artifacts suddenly resurface on the spectrum until everything has enough time to resettle again. This can be very problematic as the re-emergence of artifacts can cause desensitization.

There is yet one more problem to be aware of. In describing some of the previous algorithms,[11] it has alluded to "pathological" inputs that can lead to erroneous estimates.

For example, in the first case of digital offset estimation, if an input has a frequency that is an integer submultiple of the sample rate, each sub-ADC samples a different constant data series or periodic sequences with different dc levels and/or different periods. Even if the input is not *exactly* at one of these frequencies, but close enough to them, it is possible to observe very slowly varying data series at the outputs of each sub-ADC (at beat frequency). Because of that, the mismatch estimation converges very slowly or possibly oscillates. Meanwhile, the actual mismatches could drift faster than the convergence rate.

These are examples of a broader family of problematic signals and more accurate definitions are in order.

11. Which now can be recognized as employing blind identification.

A signal $u[k]$ is called QS if both:

$$m_u = \lim_{M\to\infty} \frac{1}{M}\sum_{k=1}^{M} E(u[k]) \tag{11a}$$

and:

$$R_u[n] = \lim_{M\to\infty} \frac{1}{M}\sum_{n=1}^{M} E(u[k+n]u[k]) \tag{11b}$$

exist.

While $u[k]$ is called *modulo-N QS* if there exists a function $g(\cdot,\cdot,\ldots)$ for which

$$\begin{aligned} g_{ui1,ui2,\ldots} &= \lim_{M\to\infty} \frac{1}{M}\sum_{k=1}^{M} g(ui1[k], ui2[k], \ldots) \\ & \quad i1, i2, \ldots = 0, \ldots, N-1 \\ R_u[n] &= \lim_{M\to\infty} \frac{1}{M}\sum_{n=1}^{M} E(u[k+n]u[k]) \end{aligned} \tag{12}$$

exist and

$$g_{ui1,ui2,\ldots} = g_{(i1+l) \bmod N,(i2+l) \bmod N,\ldots} \forall l \in \mathbb{Z} \tag{13}$$

where, in this context, $ui1[k], ui2[k], \ldots$ are the sub-ADC outputs.

Any signal that does *not* meet the latter requirements is called *nonmodulo-N* QS and it is going to cause problems for the time-interleaved blind identification algorithms since it will lead to contradictory, perhaps oscillating or diverging estimation dependent on the observation of the N sub-ADCs's outputs.

The input tones at frequency $f_{o,k}$, in the case of the offset estimation, are nonmodulo N QS. In this case, each sub-ADC returns an output data series with a *different* average, hence failing to meet (12). This makes it impossible to distinguish the different sub-ADC offsets from the different dc levels generated by this oddly regular input. More examples can be thought of, though they all have very regular patterns coherent with the order of interleaving and the sample rate. The issue lies with interleaving.

In addition to introducing appropriate filters in the algorithms, as seen in previous sections, other common mitigations include adding some form of random dithering onto the inputs or in various places within the ADC to break these oddly regular patterns. Alternatively, it is necessary to augment the algorithm with some form of detection of the occurrence of one of these conditions. If one of these conditions is detected, the identification process can be stopped until this singular state is overcome. Once favorable conditions are restored, the estimation process is resumed.

G. RANDOMIZATION

Calibration reduces the power of the TI artifacts, though complete removal may not always be possible or practical. To further reduce the TI artifacts' peak power, improving SFDR, additional mitigation is possible.

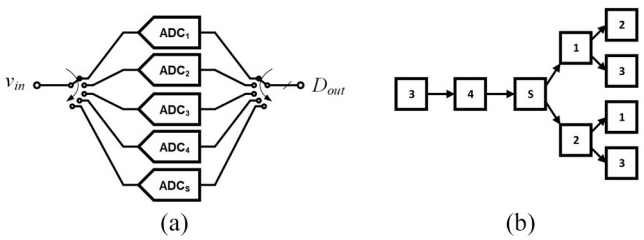

FIGURE 21. Shuffling. (a) Extended array. (b) Modified TI sequence [21], [62], [71].

Since the modulation that gives rise to the artifacts originates from mixing the input with a *periodic* error sequence, breaking this periodicity can mitigate the problem. The mismatch values do not change, but the order in which these mix with the sampled input data does. For that, it is necessary to change the order in which the mismatched functional blocks are used during the array operation [21], [62], [71]. This can be done by shuffling the order of the mismatched sub-ADCs and/or the SHAs, similarly to what is done with dynamic element matching (DEM) in DACs [35], [88].

With traditional TI ADCs, at any given time there is always *only one* sub-ADC (or SHA) that is ready to process the next sample. That locks the sampling order. To change the next sub-ADC with another one and break the sequence's order, at least one more sub-ADC (or SHA) beyond the N of the array is required. With, at least $N + 1$ sub-ADCs, it is possible to reorder the sampling sequence dynamically and to keep track of the new processing order to correctly demultiplex data back at the output. To be clear, this is still an N-times interleaved ADC since the input is still sampled by a set of N sub-ADCs. What is changing is that there is now a larger ($\geq N+1$) set of available sub-ADCs allowing order randomization.

An example is shown in Fig. 21. A four TI ADC is augmented by a spare sub-ADC, ADC$_S$, identical to the original four sub-ADCs. Once the spare ADC$_S$ is weaved in the sequence, different alternate sampling orders become possible as shown in the state diagram of Fig. 21(a) that displays possible random interleaving sequences. Because of the randomization, the artifacts' spurious power spread over the noise floor. The total spurious power is unchanged (SNDR is unchanged) as the mismatches have not been reduced or compensated, but the SFDR improves, similarly to DEM.

H. COMPARISONS AND ADDITIONAL CONSIDERATIONS

Comparing some of the previous algorithms is nontrivial, particularly because, in practice, calibration and architecture are highly interdependent. For instance, for comparable overall specifications, in a hierarchical architecture with few front-end SHAs, timing-skew correction in the analog domain is feasible and the digital part of the algorithm may be computationally less challenging than for a direct architecture with a large array of samplers. Moreover, the nature of the input signal and the specifics of each ADC application inform the algorithm selection or *when* to run it. For instance, if the ADC is used in a spectrum analyzer or other high-end lab instruments, the slow speed of calibration's convergence is generally not an issue though minimal disturbance to the measured signal is very important. Therefore, a foreground calibration can be performed with an optimal training signal at startup and then, during normal operation, a background blind calibration algorithm tracks and corrects slow parametric drifts. While, in a RADAR or a LIDAR application, background algorithms involving the injection of PN stimuli (subsequently removed from the output) are much less dependent on non-QS inputs and foreground calibration may be unnecessary. On the other hand, ADCs used to sense the power amplifier (PA) output in digital predistortion (DPD) loops are not operating continuously [17]. So, these can be foreground calibrated during idle time intervals using architecture-specific fast-convergence algorithms relying on optimal stimuli [55], [58]. With these disclaimers in mind, qualitative comparisons are summarized in Table 3. Ultimately, as in many other analog problems, design requires carefully evaluating multiple tradeoffs, possibly iterating, and reconsidering initial decisions as the implementation progresses.

TABLE 3. Algorithms summary.

Feature	Ref ADC [53,74,75, 77]	Window detection [76]	Blind sub-A/D to sub-A/D correlation [48, 70-72, 78, 80, 84, 85, 94]	Blind with PN injection [19, 66-68, 86]
Rate of convergence	Fast with sine, slower with WB	Fast	Medium	Slow
Analog overhead	Large	Medium	Small	Medium
Input dependence	High	Medium	Low	Low
Input distortion	Yes	No	No	No
Fit to high f_s	Yes	No	Yes	Yes
Computational load	Low	Medium	High	Medium

VI. EXAMPLES

Three examples discussed in this section illustrate the implementation of the principles and techniques covered in the previous sections.

The first example is the 8b/72GSPS ADC in a 14-nm FinFET process, for a wired/optical communication application, reported in [58]. In this application, bandwidth and power efficiency are critically important. While SFDR and SNDR are much more forgiving than, for example, in wireless applications. Therefore, recalling Section IV-A, high power efficiency is obtained by selecting a very efficient sub-ADC architecture as suggested by (4). The Figure of

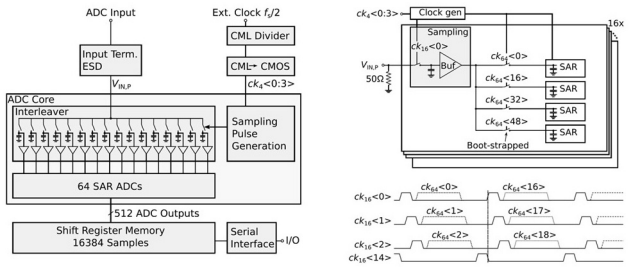

FIGURE 22. 8 b/72 GSPS 16×4 TI SAR ADC for wired communication applications in 14-nm FinFET [58].

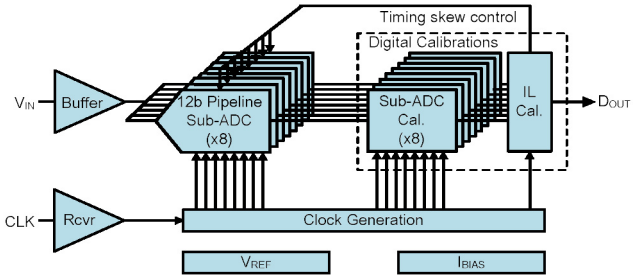

FIGURE 23. 12 b/10 GSPS 8× TI pipelined ADC for instrumentation and wireless communication applications in 28-nm CMOS [21].

Merit survey [13] shows that single SAR ADCs have excellent power efficiency for SNDR between ∼36 and ∼58 dB with little to no individual calibration, making this architecture a sensible choice for a sub-ADC in this case. SAR ADC sample rates are 1 ∼ 2GSPS, which requires high order of interleaving, i.e., N between, say, 72 and 36. A hierarchical interleaving architecture has been implemented, as in Section IV-B. The array has a first rank of $L = 16$ samplers, each one driving a second rank of $K = 4$ SAR sub-ADCs, for a total of $N = L{\cdot}K = 16{\cdot}4 = 64$ sub-ADCs, each one sampling at a maximum rate of $f_s/N = 72/64 = 1.125$ GSPS, as in Fig. 22 [58].

The input sampling bandwidth is maximized after an exhaustive assessment of the tradeoffs between sampling networks and the resulting choice on the first and second rank multiplicity L and K [58]. Calibrating the time skew between the 16 samplers is very challenging. These are controlled by adjusting the gate delays of the clock circuitry set by digitally controlled capacitive loads to the logic [58]. Offset mismatch, gain error mismatch, and sample time skew mismatch are foreground calibrated off-chip with sine wave stimuli [58]. At 72 GSPS, this ADC achieved an SNDR between 39 dB (at low input frequency) and 30 dB (at high frequency), consuming 235 mW and requiring an active area of 0.15 mm^2.

The second example is the 12 b/10 GSPS ADC in a 28-nm CMOS process intended for high-end instrumentation and wireless communication applications reported in [21]. This requires a higher dynamic range and lower signal bandwidth than the previous example: SFDR of the order of 70 dB at 1 GHz or higher and NSD of the order of −150 dBFS/Hz or better. Minimizing images' power is critically important and interleaving the minimum possible number of sub-ADCs is a practical approach [21], [53].

Pipelining, as in Section III-A, speeds up the sample rate considerably compared to the previous SAR ADC choice, especially since a much higher dynamic range is required. A direct $N = 8$ TI ADC employing pipelined sub-ADCs sampling at $f_s/N = 10/8 = 1.25$ GSPS has been used in [21], as shown in Fig. 23. The sub-ADCs are individually calibrated for a variety of sub-block nonidealities. Digital background calibration, based on the blind identification of offset mismatch, gain error mismatch, and timing skew mismatch

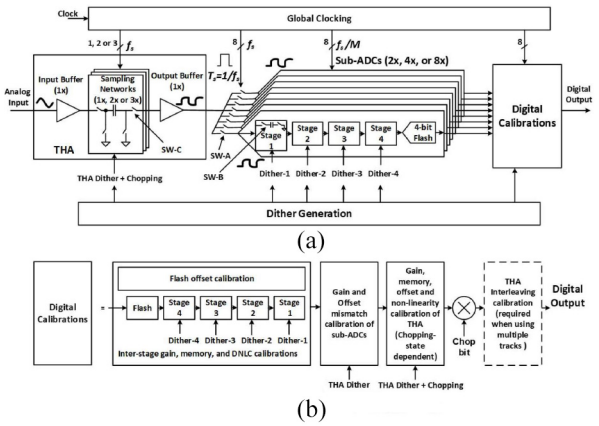

FIGURE 24. 12 b/18 GSPS 8× TI pipelined ADC in 16m FinFET. (a) Conceptual architecture. (b) Digital calibrations [20].

runs on-chip [21]. Higher SFDR is obtained by sub-ADC order shuffling. To accomplish that, the array is reconfigured for $N = 7$, and using an additional sub-ADC for order randomization as in Section V-G. An innovative buffer architecture using stacked complementary source follower drives the non-negligible front-end load [21]. At 10GSPS this ADC achieved an SNDR of 55 dB and an SFDR of 66 dB close to Nyquist, consuming 2.9 W and requiring an active area of 7.4 mm^2.

A digital post-distortion (linearization) technique allowing to substantially reduce, in-situ, the combined nonlinearity of this ADC and multiple off-chip linear stages preceding it, has been reported in [44] and [45] (∼10-dB improvement in SFDR).

The last example is the 12 b/18 GSPS ADC in a 16-nm FinFET process reported in [20] and shown in Fig. 24. Most of the previous high-level considerations apply here, leading to an analogous $N = 8$ directly interleaved pipelined sub-ADCs. There are notable differences, however, between this work and the previous one. For instance, instead of $L = N = 8$ samplers as in the previous example, in this ADC, the front-end sampling is done by a 2-times interleaved THA driving the $N = 8$ sub-ADC array. To mitigate front-end sampling mismatches, a third THA is used for shuffling. All other TI mismatches and nonidealities are dealt with multiple

TABLE 4. Ti ADC examples.

Ref.	fs, n (GSPS, bits)	Proc. Tech.	Arch. (N or LxK)	Sub-A/D	NSD (dBFS/Hz)	SFDR (dBc)	FOM (dB)
Taft [34]	1, 10	180nm	2	Pipe-FI	-143	66	143
Gupta [47]	1, 11	130nm	4	Pipe	-139	60	145
Doris [54]	2.6, 10	65nm	4x16	SAR	-143	53	142
Straayer [56]	4, 12	65nm	8x2	Pipe	-150	64	145
Le Dortz [86]	1.62, 9	40nm	12	SAR	-136	50	147
Song [76]	0.8, 10	40nm	4	SAR	-134	56	157
Devarajan [20]	10, 12	28nm	8	Pipe	-157	66	148
Ramkaj [51]	5, 12	28nm	8	Pipe-SAR	-151	65	161
Keane [52]	8, 10	28nm	16	SAR	-148	60	150
Brandolini [55]	5, 10	28nm	2x4	Pipe-SAR	-141	58	148
Guo [94]	5, 10	28nm	16	SAR	-143	60	158
Vaz [57]	4, 13	16nm	2x4	Pipe-SAR	-150	67	153
Ali [20]	18, 12	16nm	8	Pipe	-157	54	150
Kull [58]	72, 8	14nm	16x4	SAR	-106	44	146
Nguyen [59]	97, 8	7nm	16x8	SAR	-143	42	147

digital (blind) LMS loops exercised by PN sequences in various sections of the ADC [20]. At a sub-block level, digital assistance is more aggressively used for PVT compensation and for post-distortion enhancement of speed-critical blocks, allowing the use of open-loop Gm-R amplifiers [20] instead of closed-loop structures in the samplers and the pipeline sub-ADCs, as in Section III-B, achieving about twice the sample as the previous case, better FOM, smaller area, even though the process technology is not commensurately faster. At 18GSPS this ADC achieved an SNDR between 52 dB (for $f_{\text{in}} = 4$ GHz) and 48 dB (for $f_{\text{in}} = 8$ GHz) and an SFDR of 56∼54 dB, consuming 1.3 W and an active area of 2.6 mm^2.

Table 4 reports a collection of additional representative examples.

VII. CONCLUSION

To summarize, the demand for increasingly wider band digitization continues unabated by the needs of established applications, such as wired and wireless communication, high-end instrumentation, radars, and military applications, along with emerging ones, such as wide-band sensing in autonomous vehicles, smart factories, and industrial automation among others. This requires increasingly faster data converters embedded in systems on a chip (SoC) or systems on a package (SoP/SiP). This article has provided a high-level overview of some of the enabling techniques and architectures for modern high-speed Nyquist ADCs, particularly TI ADCs.

As demands evolve and technology develops, multiple challenges need solutions, motivating further research.

The ever-present challenge is managing power consumption and conversion efficiency. In some cases, high power consumption causes high die temperature, which is concerning when localized, causing "hot-spots," device ageing [89], electromigration, and other reliability concerns.

Other issues, only minimally mentioned in this article for space reasons, include the challenges of driving the input signal to the ADC [90], providing a low phase noise clock [25], ensuring that the ADC's calibration does not interfere with the system-level calibration for the signal chain embedding the ADC itself [91], enabling cost-effective manufacturability with built-in self-test [92].

From a design point of view, some of the most important open problems relate to calibration. As discussed in Sections III and IV, timing skew calibration and bandwidth mismatch calibration are very much dependent on use cases of different applications and intimately tied to architectural choices. For any calibration technique that works in a prescribed context, there are multiple ways to make it fail. This is presently a very fertile line of research where new techniques are emerging rapidly, enabling steps forward in performance. Digital assistance, architecture, circuit/transistor-level design, and physical design are inextricably tied to one another, particularly when pushing the technology limits. A solid understanding of analog design and its physical implementation is a necessary requirement that cannot be waived to resort to some generic black-box digital clean-up panacea. Comprehensive analog–digital co-design is required, and it scales in finer lithography [14].

It is the author's sincere hope that this article aids in better connecting the dots in this exciting field, and in attracting the interest of talented researchers, accelerating the solution of its open problems and pushing technology forward.

ACKNOWLEDGMENT

The author wishes to acknowledge the anonymous reviewers, the Associate Editor, Dr. Qinwen Fan, and Dr. Nevena Rakuljic of Analog Devices for their feedback, which has helped improving the quality of this manuscript, and the Editor-in-Chief, Prof. E. Cantatore for his encouragement and support.

REFERENCES

[1] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6G wireless systems: Vision, requirements, challenges, insights, and opportunities," *Proc. IEEE*, vol. 109, no. 7, pp. 1166–1199, Jul. 2021.

[2] H. Zhou, W. Xu, J. Chen, and W. Wang, "Evolutionary V2X technologies toward the Internet of Vehicles: Challenges and opportunities," *Proc. IEEE*, vol. 108, no. 2, pp. 308–323, Feb. 2020.

[3] K.-C. Chen, S.-C. Lin, J.-H. Hiao, C.-H. Liu, A. F. Molish, and G. P. Fettweis, "Wireless networked multirobot systems in smart factories," *Proc. IEEE*, vol. 109, no. 4, pp. 468–494, Apr. 2021.

[4] "Decadal plan for semiconductors." Semiconductor Research Corporation. Jan. 2021. Accessed: May 2022. [Online]. Available: https://www.src.org/about/decadal-plan/

[5] M. Bohr, "A 30 year retrospective on Dennard's MOSFET scaling paper," *IEEE Solid-State Circuits Soc. Newslett.*, vol. 12, no. 1, pp. 11–13, 2007.

[6] K. Bult, "The effect of technology scaling on power dissipation in analog circuits," in *Analog Circuit Design*. Berlin, Germany: Springer-Verlag, 2006, pp. 251–290.

[7] G. Moore, "No exponential is forever: But 'forever' can be delayed!" in *Proc. Int. Solid-State Circuits Conf.*, 2003, pp. 20–23.

[8] M. Liu, "1.1 Unleashing the future of innovation," in *Proc. Int. Solid-State Circuits Conf.*, 2021, pp. 9–16.

[9] A. Loke, "Device and physical design considerations for circuits in FinFET technology," in *Proc. Int. Solid-State Circuits Conf.*, 2020, p. 79.

[10] K. L. Loh, FinFETs for analog & mixed-signal designs' in 'forum F2: FinFETs & FDSOI—A mixed signal circuit designer's perspective,"' in *Proc. Int. Solid-State Circuits Conf.*, 2018, pp. 505–507.

[11] J. Strange, "RF, mmW and subTHz semiconductor trends & the outlook towards 6G," in *Proc. Global Foundries Tech. Conf.*, 2021.

[12] C. Hou, "1.1 A smart design paradigm for smart chips," in *Proc. Int. Solid-State Circuits Conf.*, 2017, pp. 8–13.

[13] B. Murmann. "ADC performance survey 1997–2021." Accessed: Apr. 10, 2022. [Online]. Available: http://web.stanford.edu/∼murmann/adcsurvey.html

[14] G. Manganaro, Ultra-high-data-rate ADCs and DACs: Architectures and implementations' in 'short course: High speed/high performance data converters: metrics, architectures, and emerging topics,"' in *Proc. Int. Solid-State Circuits Conf.*, 2022, pp. 567–568.

[15] G. Manganaro, *Advanced Data Converters*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[16] A. M. A. Ali, *High Speed Data Converters*. London, U.K.: IET Press, 2016.

[17] G. Manganaro and D. M. W. Leenaerts, Eds., *Advances in Analog and RF IC Design for Wireless Communication Systems*. Oxford, U.K.: Academic Press, 2013.

[18] P. Wambacq and W. Sansen, *Distortion Analysis of Analog Integrated Circuits*. Dordrecht, The Netherlands: Kluwer, 1998.

[19] M. Skolnik, *Introduction to Radar Systems*. New York, NY, USA: MCGraw-Hill, 2008.

[20] A. M. A. Ali et al., "A 12-b 18-GS/s RF sampling ADC with an integrated wideband track-and-hold amplifier and background calibration," *IEEE J. Solid-State Circuits*, vol. 55, no. 12, pp. 3210–3224, Dec. 2020.

[21] S. Devarajan et al., "A 12-b 10-GS/s interleaved pipeline ADC in 28-nm CMOS technology," *IEEE J. Solid-State Circuits*, vol. 52, no. 12, pp. 3204–3218, Dec. 2017.

[22] A. M. A. Ali et al., "A 14-bit 2.5GS/s and 5GS/s RF sampling ADC with background calibration and dither," in *Proc. IEEE Symp. VLSI Circuits*, 2016, pp. 206–207.

[23] B. Razavi, *Principles of Data Conversion System Design*. Piscataway, NJ, USA: IEEE Press, 1995.

[24] N. Da Dalt, M. Harteneck, C. Sandner, and A. Wiesbauer, "On the jitter requirements of the sampling clock for analog-to-digital converters," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 49, no. 9, pp. 1354–1360, Sep. 2002.

[25] B. Razavi, "Jitter-power trade-offs in PLLs," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 68, no. 4, pp. 1381–1387, Apr. 2021.

[26] V. J. Arkesteijn, E. A. M. Klumperink, and B. Nauta, "Jitter requirements of the sampling clock in software radio receivers," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 53, no. 2, pp. 90–94, Feb. 2006.

[27] G. Manganaro, S. U. Kwak, S. Cho, and A. Pulincherry, "A behavioral modeling approach to the design of a low jitter clock source," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 50, no. 11, pp. 804–814, Nov. 2003.

[28] U. Seng-Pan, "Hybrid-design of analog-to-digital converters: Tutorial 5," in *Proc. Int. Solid-State Circuits Conf.*, 2018, pp. 499–501.

[29] S. H. Lewis and P. R. Gray, "A pipelined 5-Msample/s 9-bit analog-to-digital converter," *IEEE J. Solid-State Circuits*, vol. 22, no. 6, pp. 954–961, Dec. 1987.

[30] J. Li, R. Leboeuf, M. Courcy, and G. Manganaro, "A 1.8V 10b 210MS/s CMOS pipelined ADC featuring 86dB SFDR without calibration," in *Proc. IEEE Custom Integr. Circuits Conf.*, 2007, pp. 317–320.

[31] B.-G. Lee, B.-M. Min, G. Manganaro, and J. W. Valvano, "A 14-b 100-MS/s pipelined ADC with a merged SHA and first MDAC," *IEEE J. Solid-State Circuits*, vol. 43, no. 12, pp. 2613–2619, Dec. 2008.

[32] R. J. Van De Plassche and R. E. J. Van Der Grift, "A high-speed 7 bit A/D converter," *IEEE J. Solid-State Circuits*, vol. JSSC-14, no. 6, pp. 938–943, Dec. 1979.

[33] R. C. Taft and M. R. Tursi, "A 100-MS/s 8-b CMOS subranging ADC with sustained parametric performance from 3.8 V down to 2.2 V," *IEEE J. Solid-State Circuits*, vol. 36, no. 3, pp. 331–338, Mar. 2001.

[34] R. C. Taft et al., "A 1.8 V 1.0 GS/s 10b self-calibrating unified-folding-interpolating ADC with 9.1 ENOB at Nyquist frequency," *IEEE J. Solid-State Circuits*, vol. 44, no. 12, pp. 3294–3304, Dec. 2009.

[35] S. Pavan, R. Schreier, and G. C. Temes, *Understanding Delta-Sigma Data Converters*, 2nd ed. Piscataway, NJ, USA: IEEE Press, 2017.

[36] C. Wu and J. Yuan, "A 12-bit, 300-MS/s single-channel pipelined-SAR ADC with an open-loop MDAC," *IEEE J. Solid-State Circuits*, vol. 54, no. 5, pp. 1446–1454, May 2019.

[37] H. Shibata et al., "A 9-GS/s 1.125-GHz BW oversampling continuous-time pipeline ADC achieving −164-dBFS/Hz NSD," *IEEE J. Solid-State Circuits*, vol. 52, no. 12, pp. 3219–3234, Dec. 2017.

[38] J. Liu, M. Hassanpourghadi, and M. S.-W. Chen, "A 10GS/s 8b 25fJ/c-s 2850um^2 two-step time-domain ADC using delay-tracking pipelined-SAR TDC with 500fs time step in 14nm CMOS technology," in *Proc. Int. Solid-State Circuits Conf.*, 2022, pp. 160–161.

[39] F.-J. Wang, G. C. Temes, and S. Law, "A quasi-passive CMOS pipeline D/A converter," *IEEE J. Solid-State Circuits*, vol. 24, no. 12, pp. 1752–1755, Dec. 1989.

[40] G. Manganaro, S.-U. Kwak, and A. R. Bugeja, "A dual 10-b 200-MSPS pipelined D/A converter with DLL-based clock synthesizer," *IEEE J. Solid-State Circuits*, vol. 39, no. 11, pp. 1829–1838, Nov. 2004.

[41] X. Jiang, Ed., *Digitally-Assisted Analog and Analog-Assisted Digital IC Design*. Cambridge, U.K.: Cambridge Univ. Press, 2015.

[42] C. Jabbour, P. Desgreys, and D. Dallet, Eds., *Digitally-Enhanced Mixed-Signal Systems*. London, U.K.: IET Press, 2019.

[43] B. Murmann and B. E. Boser, "A 12-bit 75-MS/s pipelined ADC using open-loop residue amplification," *IEEE J. Solid-State Circuits*, vol. 38, no. 12, pp. 2040–2050, Dec. 2003.

[44] N. Rakuljic, C. Speir, E. Otte, J. Bray, C. Petersen, and G. Manganaro, "In-situ nonlinear calibration of a RF signal chain," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, 2018, pp. 1–5.

[45] G. Manganaro and N. Rakuljic, "Receiver signal chains with low power drivers and driver optimization," U.S. Patent 20 200 304 135 A1, 2020.

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

[47] S. K. Gupta, M. A. Inerfield, and J. Wang, "A 1-GS/s 11-bit ADC with 55-dB SNDR, 250-mW power realized by a high bandwidth scalable time-interleaved architecture," *IEEE J. Solid-State Circuits*, vol. 41, no. 12, pp. 2650–2657, Dec. 2006.

[48] B. Razavi, "Design considerations for interleaved ADCs," *IEEE J. Solid-State Circuits*, vol. 48, no. 8, pp. 1806–1817, Aug. 2013.

[49] B. Murmann, "Energy limits in A/D converters," in *Proc. IEEE Faible Tension Faible Consommation (FTFC)*, Paris, France, Jun. 2013, pp. 1–4.

[50] A. T. Ramkaj, "Multi-GHz bandwidth power-efficient Nyquist A/D conversion," Ph.D. dissertation, Dept. Arenberg Doctoral School, Faculty Eng. Sci., Katholieke Universiteit Leuven, Leuven, Belgium, Mar. 2021.

[51] A. Ramkaj et al., "A 5-GS/s 158.6-mW 9.4-ENOB passive-sampling time-interleaved three-stage pipelined-SAR ADC with analog–digital corrections in 28-nm CMOS," *IEEE J. Solid-State Circuits*, vol. 55, no. 6, pp. 1553–1564, Jun. 2020.

[52] J. P. Keane et al., "16.5 An 8GS/s time-interleaved SAR ADC with unresolved decision detection achieving −58dBFS noise and 4GHz bandwidth in 28nm CMOS," in *Proc. Int. Solid-State Circuits Conf.*, 2017, pp. 284–285.

[53] A. Buchwald, "Practical considerations for application specific time interleaved ADCs," in *Proc. IEEE Custom Integr. Circuits Conf. (CICC)*, 2015, pp. 1–8.

[54] K. Doris, E. Janssen, C. Nani, A. Zanikopoulos, and G. Van der Weide, "A 480mW 2.6GS/s 10b 65nm CMOS time-interleaved ADC with 48.5dB SNDR up to Nyquist," in *Proc. Int. Solid-State Circuits Conf.*, 2011, pp. 180–182.

[55] M. Brandolini et al., "A 5 GS/s 150 mW 10 b SHA-less pipelined/SAR hybrid ADC for direct-sampling systems in 28 nm CMOS," *IEEE J. Solid-State Circuits*, vol. 50, no. 12, pp. 2922–2934, Dec. 2015.

[56] M. Straayer et al., "27.5 A 4GS/s time-interleaved RF ADC in 65nm CMOS with 4GHz input bandwidth," in *Proc. Int. Solid-State Circuits Conf.*, 2016, pp. 464–465.

[57] B. Vaz et al., "16.1 A 13b 4GS/s digitally assisted dynamic 3-stage asynchronous pipelined-SAR ADC," in *Proc. Int. Solid-State Circuits Conf.*, 2017, pp. 276–277.

[58] L. Kull et al., "A 24–72-GS/s 8-b time-interleaved SAR ADC with 20–3.3-pJ/conversion and >30 dB SNDR at Nyquist in 14-nm CMOS FinFET," *IEEE J. Solid-State Circuits*, vol. 53, no. 12, pp. 3508–3516, Dec. 2018.

[59] R. L. Nguyen et al., "8.6 A highly reconfigurable 40-97GS/s DAC and ADC with 40GHz AFE bandwidth and sub-35fJ/conv-step for 400Gb/s coherent optical applications in 7nm FinFET," in *Proc. Int. State Circuits Conf.*, 2021, pp. 136–138.

[60] Y.-C. Jenq, "Digital spectra of nonuniformly sampled signals: Fundamentals and high-speed waveform digitizers," *IEEE Trans. Instrum. Meas.*, vol. 37, no. 2, pp. 245–251, Jun. 1988.

[61] N. Kurosawa, H. Kobayashi, K. Maruyama, H. Sugawara, and K. Kobayashi, "Explicit analysis of channel mismatch effects in time-interleaved ADC systems," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 48, no. 3, pp. 261–271, Mar. 2001.

[62] J. Elbornsson, F. Gustafsson, and J.-E. Eklund, "Analysis of mismatch effects in a randomly interleaved A/D converter system," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 52, no. 3, pp. 465–476, Mar. 2005.

[63] C. Vogel, "The impact of combined channel mismatch effects in time-interleaved ADCs," *IEEE Trans. Instrum. Meas.*, vol. 54, no. 1, pp. 415–427, Feb. 2005.

[64] M. El-Chammas and B. Murmann, "General analysis on the impact of phase-skew in time-interleaved ADCs," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, no. 5, pp. 902–910, May 2009.

[65] S. Ghosh and B. D. Sahoo, "Closed-form expression for the combined effect of offset, gain, timing, and bandwidth mismatch in time-interleaved ADCs using generalized sampling," *IEEE Trans. Instrum. Meas.*, vol. 70, Oct. 2021, Art. no. 2007112. [Online]. Available: https://ieeexplore.ieee.org/document/9592770

[66] D. Fu, K. C. Dyer, S. H. Lewis, and P. J. Hurst, "A digital background calibration technique for time-interleaved analog-to-digital converters," *IEEE J. Solid-State Circuits*, vol. 33, no. 12, pp. 1904–1911, Dec. 1998.

[67] S. M. Jamal, D. Fu, N. C.-J. Chang, P. J. Hurst, and S. H. Lewis, "A 10-b 120-Msample/s time-interleaved analog-to-digital converter with digital background calibration," *IEEE J. Solid-State Circuits*, vol. 37, no. 12, pp. 1618–1627, Dec. 2002.

[68] K. C. Dyer, D. Fu, S. H. Lewis, and P. J. Hurst, "An analog background calibration technique for time-interleaved analog-to-digital converters," *IEEE J. Solid-State Circuits*, vol. 33, no. 12, pp. 1912–1919, Dec. 1998.

[69] T.-H. Tsai, P. J. Hurst, and S. H. Lewis, "Bandwidth mismatch and its correction in time-interleaved analog-to-digital converters," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 53, no. 10, pp. 1133–1137, Oct. 2006.

[70] J. Elbornsson, F. Gustafsson, and J.-E. Eklund, "Blind adaptive equalization of mismatch errors in a time-interleaved A/D converter system," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 1, pp. 151–158, Jan. 2004.

[71] J. Elbornsson, F. Gustafsson, and J.-E. Eklund, "Blind equalization of time errors in a time-interleaved ADC system," *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 1413–1424, Apr. 2005.

[72] S. Huang and B. C. Levy, "Blind calibration of timing offsets for four-channel time-interleaved ADCs," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 54, no. 4, pp. 863–876, Apr. 2007.

[73] K. C. Dyer, J. P. Keane, and S. Lewis, "Calibration and dynamic matching in data converters: Part 2: Time-interleaved analog-to-digital converters and background-calibration challenges," *IEEE Solid-State Circuits Mag.*, vol. 10, no. 3, pp. 61–70, Aug. 2018.

[74] 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.

[75] A. Buchwald, "A supposedly clever thing I'll never do again," in *Proc. IEEE Custom Integr. Circuits Conf. (CICC)*, 2017, pp. 1–8.

[76] J. Song, K. Ragab, X. Tang, and N. Sun, "A 10-b 800-MS/s time-interleaved SAR ADC with fast variance-based timing-skew calibration," *IEEE J. Solid-State Circuits*, vol. 52, no. 10, pp. 2563–2575, Oct. 2017.

[77] D. Stepanović and B. Nikolić, "A 2.8 GS/s 44.6 mW time-interleaved ADC achieving 50.9 dB SNDR and 3 dB effective resolution bandwidth of 1.5 GHz in 65 nm CMOS," *IEEE J. Solid-State Circuits*, vol. 48, no. 4, pp. 971–982, Apr. 2013.

[78] C. Vogel, M. Hotz, S. Saleem, K. Hausmair, and M. Soudan, "A review on low-complexity structures and algorithms for the correction of mismatch errors in time-interleaved ADCs," in *Proc. IEEE NEWCAS*, 2012, pp. 349–352.

[79] C. Vogel and S. Mendel, "A flexible and scalable structure to compensate frequency response mismatches in time-interleaved ADCs," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, no. 11, pp. 2463–2475, Nov. 2009.

[80] S. Saleem and C. Vogel, "Adaptive blind background calibration of polynomial-represented frequency response mismatches in a two-channel time-interleaved ADC," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 58, no. 6, pp. 1300–1310, Jun. 2011.

[81] M. Seo, M. J. W. Rodwell, and U. Madhow, "Comprehensive digital correction of mismatch errors for a 400-msamples/s 80-dB SFDR time-interleaved analog-to-digital converter," *IEEE Trans. Microw. Theory Techn.*, vol. 53, no. 3, pp. 1072–1082, Mar. 2005.

[82] H. Johansson, P. Lowenborg, and K. Vengattaramane, "Least-squares and minimax design of polynomial impulse response FIR filters for reconstruction of two-periodic nonuniformly sampled signals," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 54, no. 4, pp. 877–888, Apr. 2007.

[83] S. Tertinek and C. Vogel, "Reconstruction of nonuniformly sampled bandlimited signals using a differentiator–multiplier cascade," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 55, no. 8, pp. 2273–2286, Sep. 2008.

[84] H. Johansson and P. Löwenborg, "A least-squares filter design technique for the compensation of frequency response mismatch errors in time-interleaved A/D converters," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 55, no. 6, pp. 1154–1158, Nov. 2008.

[85] H. Le Duc et al., "All-digital calibration of timing skews for TIADCs using the polyphase decomposition," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 63, no. 1, pp. 99–103, Jan. 2016.

[86] N. Le Dortz et al., "22.5 A 1.62GS/s time-interleaved SAR ADC with digital background mismatch calibration achieving interleaving spurs below 70dBFS," in *Int. Solid-State Circuits Conf. Dig.*, 2014, pp. 386–388.

[87] A. Papoulis, *Probability, Random Variable and Stochastic Processes*, 4th ed. Boston, MA, USA: McGraw-Hill, 2002.

[88] G. Manganaro, "An introduction to high data rate current-steering Nyquist DACs: Fasten your seat belts," *IEEE Solid-State Circuits Mag.*, vol. 14, no. 3, pp. 24–40, Aug. 2022.

[89] P. Ferguson and G. Manganaro, "Limiting aging effects in analog differential circuits," U.S. Patent 9 531 398 B2, 2016.

[90] A. Ramkaj, A. Cantoni, G. Manganaro, S. Devarajan, M. S. J. Steyaert, and F. Tavernier, "A 30GHz-BW <-57dB-IM3 direct RF receiver analog front end in 16nm FinFET," in *Proc. IEEE Symp. VLSI Technol. Circuits (VLSI)*, Honolulu, HI, USA, Jun. 2022, pp. 100–101.

[91] G. Manganaro, "Design considerations for direct RF sampling and synthesis wideband radios," in *Proc. Workshop WSJ IEEE Radio Freq. Integr. Circuits Symp.*, Denver, CO, USA, Jun. 2022.

[92] O. Eliezer, "Design for manufacturability in high-volume of extensively digital RF transceivers," in *Proc. Workshop WSJ IEEE Radio Freq. Integr. Circuits Symp.*, Denver, CO, USA, Jun. 2022.

[93] W. Jiang, Y. Zhu, M. Zhang, C.-H. Chan, and R. P. Martins, "A 7.6mW 1GS/s 60dB SNDR single-channel SAR-assisted pipelined ADC with temperature-compensated dynamic GM-R-based amplifier," in *Proc. Int. Solid-State Circuits Conf.*, 2019, pp. 60–62.

[94] M. Guo, J. Mao, S.-W. Sin, H. Wei, and R. P. Martins, "A 29mW 5GS/s time-interleaved SAR ADC achieving 48.5dB SNDR with fully-digital timing-skew calibration based on digital-mixing," in *Proc. Symp. VLSI Circuits*, 2019, pp. C76–C77.

GABRIELE MANGANARO (Fellow, IEEE) received the Dr.Eng. and Ph.D. degrees in electronics from the University of Catania, Catania, Italy, in 1994 and 1997, respectively.

Starting in 1994, he did research with ST Microelectronics, Catania, Italy, and Texas A&M University, College Station, TX, USA. From 1998 to 2001, he was a Member of the Technical Staff with the Data Converter Design Department, Texas Instruments Incorporated, Dallas, TX, USA. From 2001 to 2004 he was a Director of Analog Baseband Design with Engim Inc., Acton, MA, USA. In 2004 he joined National Semiconductor, holding various positions, including a Design Center Manager for the Design Center, Salem, NH, USA; and the Design Center, Munich, Germany; and eventually, as the sole Design Director for the worldwide High-Speed Data Conversion Business Unit. From 2010 to 2020 he was with Analog Devices, Wilmington, MA, USA, first as an Engineering Director for High Speed Data Converters, and, eventually, as a Technology Director of the Wireless Communication Products Division. He joined MediaTek, Woburn, MA, USA, in 2021, as a Director of Technology, presently working for the Strategic Technology Exploration Platform organization. He authored/coauthored more than 60 peer-reviewed papers, three books, notably *Advanced Data Converters* (Cambridge University Press, 2011) and has been granted 19 U.S. patents, with more pending.

Dr. Manganaro was recipient of scientific awards, including the 1995 CEU Award from the Rutherford Appleton Laboratory (U.K.), the 1999 IEEE Circuits and Systems Outstanding Young Author Award, and the 2007 IEEE European Solid-State Circuits Conference Best Paper Award. He served for the Technical Sub-Committee for Data Converters of the ISSCC for seven consecutive years. He has received international technical and leadership recognition. He is often invited to speak on technical subjects at key conferences and top universities and developed and taught graduate professional courses in data converters design and mixed-signal design with the University of Oxford, Oxford, U.K. He was an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II: EXPRESS BRIEFS from 2006 to 2007 and then an Associate Editor from 2008 to 2009, the Deputy Editor-in-Chief from 2010 to 2011, and finally the Editor-in-Chief from 2012 to 2013 for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I: REGULAR PAPERS. He is presently the Editor-in-Chief from 2020 to 2023 for the IEEE OPEN JOURNAL OF CIRCUITS AND SYSTEMS. He is a Fellow of the IET since 2009, a member of the scientific honor society Sigma Xi, and was a member of the Board of Governors for the IEEE Circuits and Systems Society from 2016 to 2018, a Distinguished Lecturer for the IEEE Solid-State Circuits Society from 2019 to 2020, he has been an Industry Advisory Member for *Proceedings of the IEEE* since 2019 and a member of multiple technical committees of the Semiconductor Research Corporation since 2019.