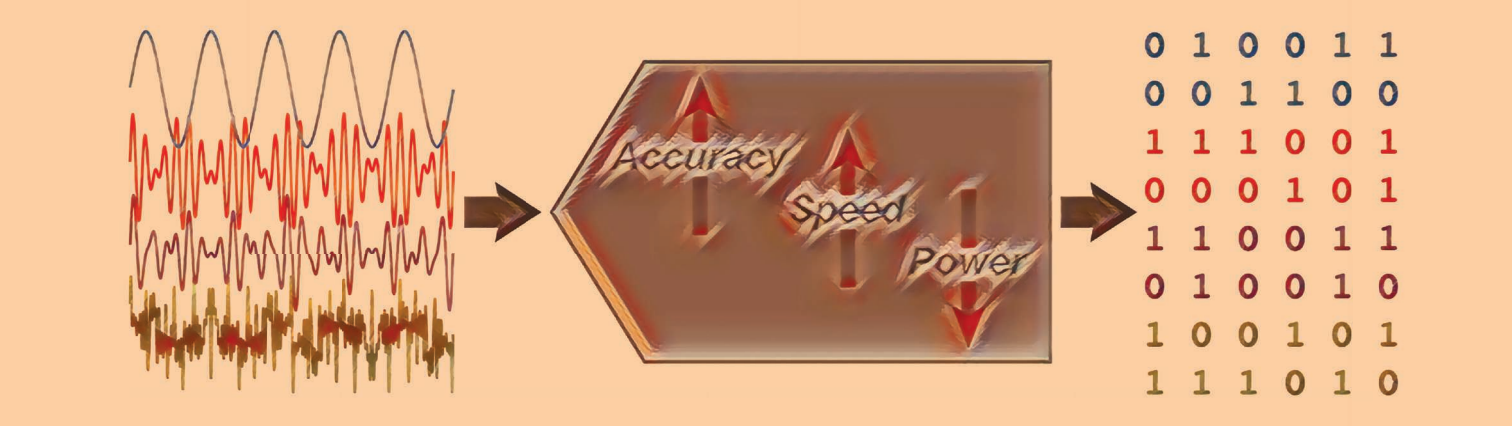
Athanasios Ramkaj, Marcel J.M. Pelgrom, Michiel S.J. Steyaert, and Filip Tavernier

In the Pursuit of the Optimal Accuracy–Speed–Power Analog-to-Digital Converter Architecture

A mathematical framework

The everlasting challenge in the design of every analog-to-digital converter (ADC) lies in maximizing the *accuracy·speed÷power* product by pushing all metrics toward their desired directions. To this end, tremendous progress has been made in advancing ADC performance both circuit- and architecture-wise. These advances have been captured by means of comparing experimental data points in surveys, with [1] being the most noteworthy. However, such comparisons give an ill-defined view since the data points correspond to different architectures that were optimized under different constraints and implemented in different process nodes. This provides little insight on architectural limits and makes a direct comparison under similar assumptions nontrivial.

This article introduces a mathematical framework to systematically estimate and compare the accuracy–speed–power limits of different ADC architectures with a complete decomposition of the blocks' contributions. (Speed refers to both sample rate and bandwidth in the sense that they are tightly coupled,



Digital Object Identifier 10.1109/MSSC.2021.3128310
Date of current version: 24 January 2022

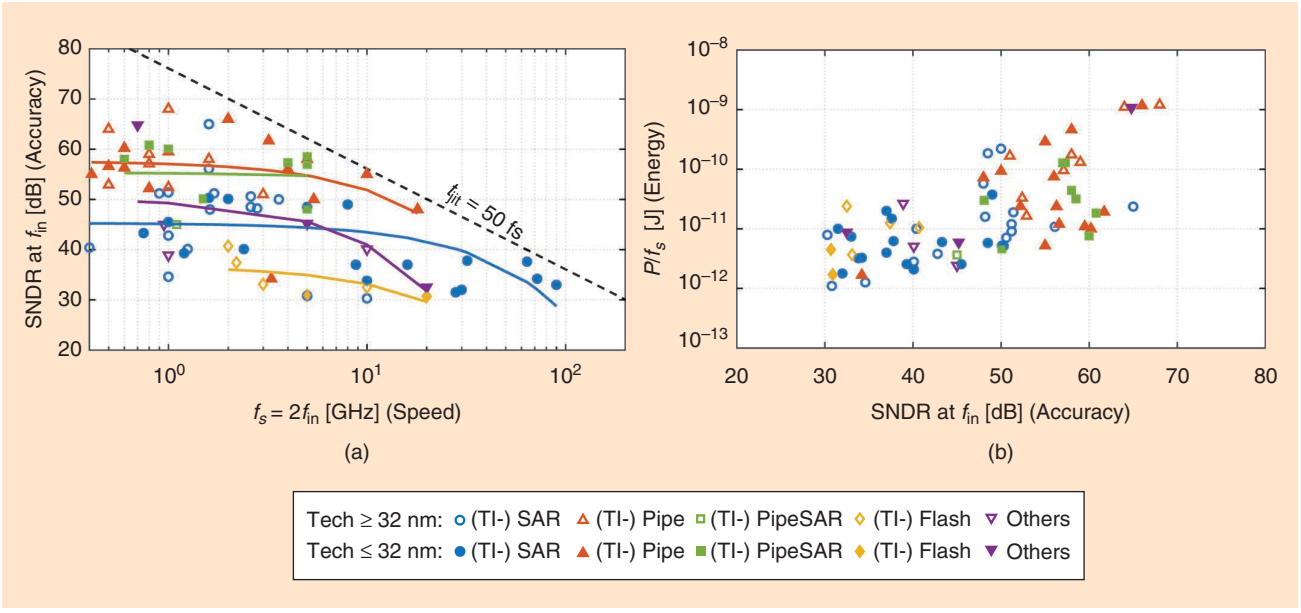


FIGURE 1: State-of-the-art design of various ADC architectures (data from [1]): (a) accuracy–speed, and (b) accuracy–energy.

whereas accuracy encapsulates both aggregate and effective resolution.) The fundamental limits of elementary building blocks, some of which have been analyzed by [2] and [3], and form the basis of our framework, are enhanced by architectural overheads and process effects, and serve as a valuable guideline when starting a new design and/or changing process nodes. The presented analysis collectively captures architectures, variants, and contributions not shown in previous literature and includes process effects beyond the typical supply scaling.

Our investigation starts by examining the state-of-the-art standings with data from [1], shown in Figure 1. Flash, successive approximation register (SAR), pipeline, pipelined-SAR, and other (ΣΔ, time-based) architectures are included, with their time-interleaved counterparts, and an older/newer-than 32-nm CMOS process node distinction. One can notice that the architectures that are most widely used across a wide range of specifications are the SAR and the pipeline, while there is a recent emergence of the pipelined-SAR. A comparison of the accuracy–speed data points after first-order curve fitting [Figure 1(a)], places the pipeline as the winner for combined high accuracy and speed. The time-interleaved–SAR is the leader in absolute speed for medium accuracy, while its speed has benefited more than the pipeline from technology scaling. The pipelined-SAR is on par with the pipeline in accuracy while approaching similar speed levels. The flash mainly resides at lower accuracies owing to its hardware-intensive nature. The accuracy–energy data points [Figure 1(b)] reveal that the SAR is leading in efficiency for medium-to-high accuracies, while the pipeline can achieve higher absolute accuracy levels. The efficiency of both architectures has improved with technology scaling, while the SAR shows an increasing trend of moving to lower accuracy and higher speed. The pipelined-SAR demonstrates the efficiency of the SAR while approaching the accuracy levels of the pipeline. This hybrid benefits from technology scaling as well since most of the designs are implemented in advanced-process nodes. The flash shows an efficiency that is similar to that of the SAR, that is, in the low-accuracy regime.

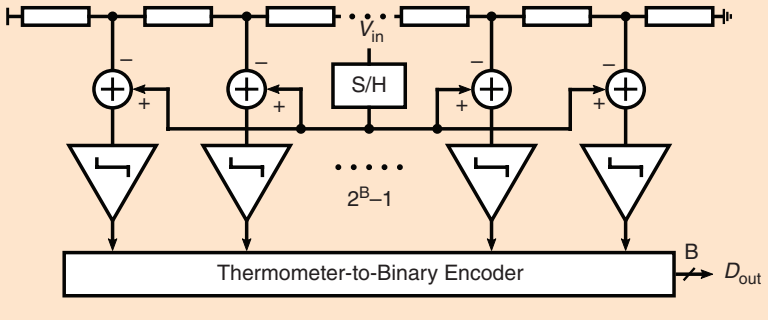


FIGURE 2: A *B*-bit flash ADC (the S/H is optional).

It is worth noting that these early conclusions are a mere interpretation of the published data, which, as already stressed, does not provide a complete picture of architectural limits and capabilities. The introduced framework is an attempt to shed some light on how fundamentally sound this interpretation may be. To this end, the key architectures highlighted above are first briefly reviewed prior to deriving their accuracy–speed–power limits.

Flash Accuracy–Speed–Power Limits

The flash ADC (Figure 2) relies on a full parallelism to quantize a signal in a single-clock cycle [4], [5]. For B-bits, it comprises 2^B-1 comparators to simultaneously compare the input signal with 2^B-1 equally spaced reference levels, given by a resistive ladder. Each comparator evaluates the difference between the input and its corresponding reference tap, outputting a thermometer code of digital 0s and 1s. A thermometer-to-binary encoder translates the resulting thermometer word into the final binary format at the ADC output. Because of their parallelism, flash ADCs achieve the highest speed among single-channel architectures. Their speed is limited by the comparator and encoder delays. The comparator regenerates exponentially on an input [3] with a time constant $\tau = C_L/g_m$, with C_L being the comparator output load and g_m the latch transconductance. τ depends on the process cutoff frequency f_T, which typically increases with scaling to finer-process nodes [6]. Thus, the flash architecture theoretically benefits from scaling, as seen in Figure 1(a). Further, owing to its minimal latency, flash is a good candidate for feedback systems.

However, the exponential comparator increase with B sets an upper bound on resolution. For each added bit, the comparator number doubles while requiring twice the accuracy. For a noise-limited design, this leads to a 4×-power for each comparator and a total of 8×-power for the ADC. Aside from that, the increasing nonlinear capacitive load to the input results in bandwidth loss and signal-dependent distortion. To date, full flash converters are limited to 8 bits. Another issue is the multicomparator kickback, creating glitches at the input and reference taps, which, because of the different impedances seen, may lead to evaluation errors. Finally, the difference in the input referred offset [7] among the comparators results in additional nonlinearity.

The flash accuracy–speed–power limits' derivation starts with its total power consumption, including sampler, comparator, ladder, and digital logic (encoder) contributions

$$P_{F,\mathrm{tot}} = P_{F,\mathrm{samp}} + (2^B-1)\cdot P_{F,\mathrm{comp}} + P_{F,\mathrm{lad}} + P_{F,\mathrm{dig}}. \quad (1)$$

$P_{F,\mathrm{samp}}$ is given by the following generalized expression:

$$P_{F,\mathrm{samp}} = V_{\mathrm{DD}}\cdot NTC\cdot \vartheta_{F,\mathrm{samp}}\cdot f_s\cdot C_S \cdot(\varphi_{\mathrm{sup}} V_{\mathrm{DD}}). \quad (2)$$

C_S is the sampling capacitance, NTC depicts the number of settling time constants, $\vartheta_{F,\mathrm{samp}}$ captures the portion of the total period $1/f_s$ allocated to sampling, and φ_{sup} is a supply use percentage ($V_{\mathrm{sig}}/V_{\mathrm{DD}}$). In the same way, $P_{F,\mathrm{comp}}$, modeling the comparator as a two-stage integrator-latch, is found as

$$P_{F,\mathrm{comp}} = V_{\mathrm{DD}}\cdot \vartheta_{F,\mathrm{comp}}\cdot f_s\cdot\Big[\{C_I\cdot\Delta V_I\} + \Big\{[(B-\log_2 A_I)\cdot\ln 2 + \ln \mathrm{BER}^{-1}]\cdot C_L\cdot\frac{V_{GT}}{2}\Big\}\Big]. \quad (3)$$

The first pair of the curly brackets includes the contribution of the noise-critical input integrator with A_I-gain, while the second pair contains the metastability-critical latch contribution [via BER] from its input/output exponential regeneration $V_{\mathrm{out}} = A_I\cdot V_{\mathrm{in}}\cdot\exp\{g_{m,L}/(\vartheta_{\mathrm{comp}}\cdot f_s\cdot C_L)\}$, and the basic MOS equation linking $g_{m,L}, I_L, V_{GT}$ [6]. $\vartheta_{F,\mathrm{comp}}$ captures the $1/f_s$-period fraction occupied by the comparator. Allocating a portion of a total noise budget to the above expressions, [3] determines C_S, C_I, and C_L, leading to fundamental limits imposed by physical quantities. In fact, the process used puts a lower limit to the value of these capacitances through minimum realizable gate sizing, which we denote as C_{min}. Further, the comparator input capacitive load, $C_{\mathrm{in},I}$, to the sampler imposes a considerable power overhead. $C_{\mathrm{in},I}$ is related to the process, f_T; the speed, f_s; and C_I via the following expressions:

$$f_T \approx \frac{g_{m,I}}{2\pi C_{\mathrm{in},I}}, \quad (4a)$$

$$f_{\mathrm{bw},I} \approx \begin{cases}\dfrac{g_{m,I}}{2\pi A_I C_I}\\ \dfrac{\vartheta_{F,\mathrm{comp}}\cdot f_s}{2\pi}\end{cases}, \quad (4b)$$

$$C_{\mathrm{in},I} \approx A_I C_I\cdot\Big(\frac{\vartheta_{F,\mathrm{comp}}\cdot f_s}{\pi f_T}\Big), \quad (5)$$

where (5) results from dividing (4*a*) by (4*b*), $f_{\mathrm{bw},I}$ denotes the comparator input integrator bandwidth (100% slewing), and a 2× overestimation is included to capture practical overhead (e.g., interconnect). Similarly, the parasitic loading at the output of each block is considered, taking into account its process, f_T, for a given biasing (V_{GT}), the parasitic loading at the output of each block is considered

$$C_{I,\mathrm{tot}} = C_I + \frac{g_{m,I}\cdot C_{I,\mathrm{par}}}{g_{m,I}} \approx C_I + \frac{g_{m,I}}{\pi f_T}, \quad (6a)$$

$$C_{L,\mathrm{tot}} = C_L + \frac{g_{m,L}\cdot C_{L,\mathrm{par}}}{g_{m,L}} \approx C_L + \frac{g_{m,L}}{\pi f_T}. \quad (6b)$$

Incorporating all the above to (2) and (3), $P_{F,\mathrm{samp}}$ and $P_{F,\mathrm{comp}}$ assume their complete forms as in (7) and (8) shown at the bottom of this page

$$P_{F,\mathrm{samp}} = V_{\mathrm{DD}}\cdot NTC\cdot\vartheta_{F,\mathrm{samp}}\cdot f_s\cdot\Big[\max\{C_S, C_{\mathrm{min}}\} + (2^B-1)\cdot\max\Big\{A_I C_I\cdot\Big(\frac{\vartheta_{F,\mathrm{comp}}\cdot f_s}{\pi f_T}\Big), C_{\mathrm{min}}\Big\}\Big]\cdot(\varphi_{\mathrm{sup}} V_{\mathrm{DD}}). \quad (7)$$

$$P_{F,\mathrm{comp}} = V_{\mathrm{DD}}\cdot\vartheta_{F,\mathrm{comp}}\cdot f_s\cdot\left[\left\{\frac{\max\{C_I, C_{\mathrm{min}}\}\cdot\Delta V_I}{1-\dfrac{2\Delta V_I}{\pi V_{GT}}\cdot\dfrac{\vartheta_{F,\mathrm{comp}}\cdot f_s}{f_T}}\right\} + \left\{\frac{[(B-\log_2 A_I)\cdot\ln 2 + \ln \mathrm{BER}^{-1}]\cdot\max\{C_L, C_{\mathrm{min}}\}\cdot\dfrac{V_{GT}}{2}}{1-\dfrac{[(B-\log_2 A_I)\cdot\ln 2 + \ln \mathrm{BER}^{-1}]}{\pi}\cdot\dfrac{\vartheta_{F,\mathrm{comp}}\cdot f_s}{f_T}}\right\}\right]. \quad (8)$$

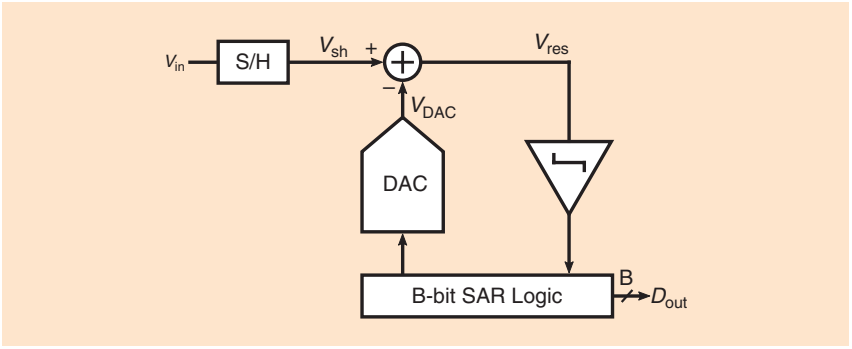


FIGURE 3: A *B*-bit SAR ADC.

$P_{F,\text{lad}}$ is derived by considering the static power through the ladder with a resistance sized for a portion of the total noise budget

$$P_{F,\text{lad}} = V_{DD}\cdot\left(\frac{\varphi_{\sup}V_{DD}}{R_{\text{lad}}}\right) = V_{DD}\cdot\left(\frac{4kT\cdot\vartheta_{F,\text{samp}}\cdot f_s}{\overline{V_{R_{\text{lad}}}^2}}\right)\cdot(\varphi_{\sup}V_{DD}), \quad (9)$$

where it is assumed that the integration noise bandwidth is equal to the $1/f_s$ fraction allocated to the sampler. Finally, $P_{F,\text{dig}}$ is computed by estimating the number of minimum-sized gates necessary for the encoder. Assuming a Wallace encoder as in [2] running at f_s, we get the following:

$$P_{F,\text{dig}} = V_{DD}\cdot f_s\cdot 5\cdot(2^B - B)\cdot C_{\min}\cdot V_{DD}. \quad (10)$$

SAR Accuracy–Speed–Power Limits

The SAR ADC relies on a bit-at-a-cycle approximation of the input signal [8], [9]. Its basic components (Figure 3) are the S/H to sample the analog input, the DAC to give binary-weighted references, and a comparator to evaluate the difference between the sampled input and the DAC output. Different topologies exist for the DAC, with the CDAC currently prevailing owing to its merging ability of the S/H and DAC functions, its dynamic nature, and its superior linearity over its resistive or current-steering equivalents. Several switching schemes exist [10], which are not detailed here for the sake of brevity. The SAR logic collects the bits and stores the comparator decision, based on which it generates the next reference to minimize the difference at the summing node, V_{res}. During conversion, the full range is divided by 2 in every cycle by appropriate grouping of the DAC elements. After sufficient DAC settling, the comparator evaluates the polarity between V_{sh} and V_{DAC} and outputs the first bit. The procedure continues until all bits are evaluated, and the *B*-bit output is collected. In contrast to the flash, the SAR resolves *B* bits in *B* cycles, but with minimum hardware. The SA algorithm is not limited to binary weights, and the input can be still approximated within a given accuracy with nonbinary (<2) weights, but with extra cycles.

Despite the remarkable efficiency of the SAR for low-medium resolution and megahertz speeds [Figure 1(b)], its bit/cycle nature remains the main bottleneck in greatly extending the speed of this converter while preserving its efficiency. Several speed-boosting techniques have been proposed to tackle this bottleneck [9] but are out of the scope of this article. Further, when the resolution increases above 10 bits, driving the noise-limited input capacitance at gigahertz speeds with low noise and distortion, high-energy efficiency becomes extremely challenging.

The SAR accuracy–speed–power limits are derived by estimating the power consumption of a binary-weighted synchronous converter, including sampler, comparator, CDAC, and SA logic contributions

$$P_{S,\text{tot}} = P_{S,\text{samp}} + B\cdot P_{S,\text{comp}} + P_{S,\text{CDAC}} + P_{S,\text{dig}}. \quad (11)$$

Incorporating the parasitic loading and process effects described in the previous section, "SAR Accuracy–Speed–Power Limits," $P_{S,\text{samp}}$ and $P_{S,\text{comp}}$ take their final forms as follows:

$$P_{S,\text{samp}} = V_{DD}\cdot NTC\cdot\vartheta_{S,\text{samp}}\cdot f_s\cdot\left[\max\{C_S, C_{\min}\} + \max\left\{A_I C_I\cdot\left(\frac{\vartheta_{S,\text{comp}}\cdot f_s}{\pi f_T}\right), C_{\min}\right\}\right]\cdot(\varphi_{\sup}V_{DD}), \quad (12)$$

and in (13), at the bottom of this page, where $\vartheta_{S,\text{samp}} = B+1$ and $\vartheta_{S,\text{comp}} = (B+1)\cdot 1/\beta, 0.7\geq\beta\geq 0.5$ capture the $1/f_s$-period fraction given to the sampler and comparator, assuming a synchronous ADC with smart comparator-CDAC time-sharing [9].

For $P_{S,\text{CDAC}}$, a symmetrical energy-saving switching is assumed [9], [10]. The reference, V_{REF}, is provided by a regulating circuit drawing current from V_{DD} with an efficiency factor, λ_{ref}, leading to

$$P_{S,\text{CDAC}} = \frac{V_{DD}}{\lambda_{\text{ref}}}\cdot\vartheta_{S,\text{CDAC}}\cdot f_s\cdot\sum_{j=1}^{B-1}2^{B-3-2j}\cdot(2^j-1)\cdot\max\left\{\frac{C_S}{2^B}, C_{\min}\right\}\cdot V_{\text{REF}}. \quad (14)$$

The above gives the average switching over a full ADC period, and $\vartheta_{S,\text{CDAC}} = [(B+1)/B]\cdot 1/(1-\beta)$ considers the comparator–CDAC time-sharing. A value of 60% is assumed

$$P_{S,\text{comp}} = V_{DD}\cdot\vartheta_{S,\text{comp}}\cdot f_s\cdot\left[\left\{\frac{\max\{C_I, C_{\min}\}\cdot\Delta V_I}{1-\frac{2\Delta V_I}{\pi V_{GT}}\cdot\frac{\vartheta_{S,\text{comp}}\cdot f_s}{f_T}}\right\} + \left\{\frac{[(B-\log_2 A_I)\cdot\ln 2 + \ln \text{BER}^{-1}]\cdot\max\{C_L, C_{\min}\}\cdot\frac{V_{GT}}{2}}{1-\frac{[(B-\log_2 A_I)\cdot\ln 2+\ln\text{BER}^{-1}]}{\pi}\cdot\frac{\vartheta_{S,\text{comp}}\cdot f_s}{f_T}}\right\}\right], \quad (13)$$

for λ_{ref}, which is reasonable for a class-AB circuit. To give a baseline for $P_{S,dig}$, of each bit, a total of five gates is assumed, three for control, one for clocking, and one for storing, leading to

$$P_{S,dig} = V_{DD} \cdot f_s \cdot 5 \cdot B \cdot C_{min} \cdot V_{DD}. \quad (15)$$

Pipeline Accuracy–Speed–Power Limits

One of the most popular architectures to realize high resolution and high speed is the pipeline ADC [11], [12], (Figure 4). It incorporates a cascade of m-stages with B_s-bits each ($s = 1 \ldots m < B$), containing an S/H, a flash sub-ADC, a sub-DAC, and a residue-amplifier (RA) with A_s-gain to amplify the difference at the summing node, V_{res}. The sub-DAC and RA combination, known as the multiplying-DAC, provides the input to the next stage. A_s may be allocated a value, 2^{B_s}, which conveniently matches the range of each stage with its back end but does not allow any margin to absorb sub-ADC errors. Redundancy is often incorporated by making A_s smaller than 2^{B_s}, with 2^{B_s-1} being a common value, and using the remaining range for error absorption. Consecutive stages operate in opposite clock phases, and the overall pipeline outputs a total $B = B_1 + B_2 + \cdots + B_m$ bits. The digital logic combines the bits in a weighted sum fashion and generates a new output with a latency of m clock periods. Pipelining allows for an increased throughput bound by the speed of only one stage, enabling very high speed. Unlike flash ADCs, there is a linear, rather than exponential, increase in hardware with B.

Two key design considerations are stage scaling and B_s. Although many options exist, an optimum scaling factor of roughly $1/A_s$ is common, and its supporting analysis is presented by [13]. Regarding B_s, reducing it allows for a faster multiplying-DAC settling and a lower flash sub-ADC loading, but this comes at the expense of more pipeline stages. On the contrary, increasing B_s reduces the stage count and relaxes the precision requirements on each stage RA; however, this leads to an increased flash sub-ADC loading, inducing a power burden when pushing the speed.

Although more hardware efficient than the flash and faster than the SAR, the pipeline's main limitation is the requirement for accurate RA, with stringent speed, noise, linearity, and power requirements. Open-loop amplifiers [12] or integrators [14] have been adopted, resulting in power savings, given that the overhead of the necessary error correction does not overcome these savings. The pipeline latency also poses an issue in feedback systems.

The pipeline accuracy–speed–power limits start by considering B_s-bit stages including 2×-redundancy, resulting in B_s–1 effective bits ($B_s - 1 = 1, 2, 3, 4$). The optimum scaling factor $1/A_s$ is assumed, however, stopping at the process C_{min}, and the m-stages are determined by adding $\leq B_s$ bits to the last stage to realize the total resolution, B. An S/H-less input is assumed. Including sampler, RA, comparator, ladder, and logic contributions along the entire pipeline, the power of a B_s-bit/stage m-stage is generally expressed as

$$P_{P,tot} = P_{P,samp} + P_{P,RA,tot} + (2^{B_s} - 1) \cdot P_{P,comp,tot} + m \cdot (P_{P,lad} + P_{P,dig}). \quad (16)$$

For $P_{P,RA,tot}$, the basic open-loop $g_{m,RA}$-C_{RA} amplifier is considered [2], with a gain of 2^{B_s-1} and purely linear settling to ¼ LSB accuracy of the back end at a percentage, ζ_{set}, of the RA time. The contribution of all the RAs along the pipeline is factored by considering the C_{RA} asymptotic expansion [3], including stage scaling. The same holds for the sub-flash comparators, $C_{in,1}$, loading the RA, given by (5) with the appropriate $1/f_s$-period fraction, $\vartheta_{P,comp}$, in the pipeline. Including all loading and process effects at the RA output, $g_{m,RA}$ is found as

$$g_{m,RA,tot} = A_s \cdot \frac{\vartheta_{P,RA}}{\zeta_{set}} \cdot f_s \cdot \frac{(B-(B_s-1)+2) \cdot \ln 2 \cdot \max\{C_{RA}, C_{min}\}}{1 - A_s \cdot \frac{(B-(B_s-1)+2) \cdot \ln 2}{\pi} \cdot \frac{\vartheta_{P,RA} \cdot f_s}{\zeta_{set} \cdot f_T}} \times \left[\sum_{i=0}^{m-2} \max\{C_{RA} \cdot 2^{-(B_s-1) \cdot i}, C_{min}\} + (2^{B_s} - 1) \cdot \sum_{i=0}^{m-2} \max\{C_{in,1} \cdot 2^{-(B_s-1) \cdot i}, C_{min}\} \right]. \quad (17)$$

In the above, $\vartheta_{P,RA}$ captures the $1/f_s$-period fraction given to the RA. In a typical pipeline, the RA shares half of the period with the sub-flash (about 70% of $0.5/f_s$), and part of that time (commonly 50%) is allocated to ζ_{set}. C_{RA} is determined for a given RA noise similar to the sampler and comparator. With a final addition to the $g_{m,RA,tot}$ of a linearity factor, $\eta_{lin} \leq 1$, as in [2], to capture the overhead for a sufficient precision to drive the back end, $P_{P,RA,tot}$ is written as

$$P_{P,RA,tot} = V_{DD} \cdot I_{RA,tot} = \frac{V_{DD}}{\eta_{lin}} \cdot g_{m,RA,tot} \cdot \frac{V_{GT}}{2}. \quad (18)$$

For $P_{P,samp}$, all the previous loading and process contributions are included with the addition of the RA input loading to the sampler $C_{in,RA}$,

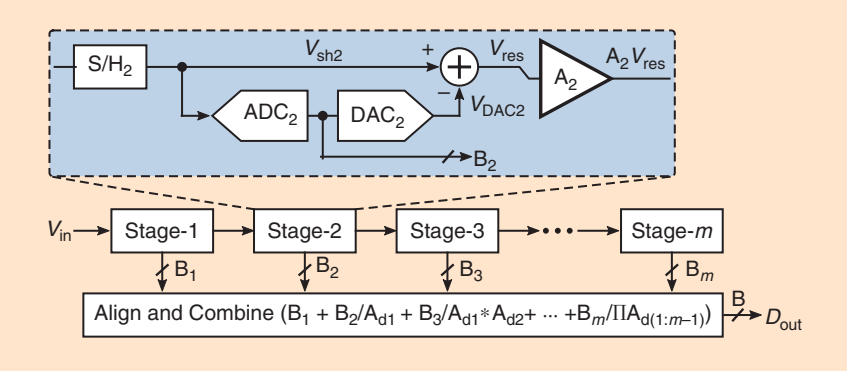


FIGURE 4: A B-bit m-stage pipeline ADC.

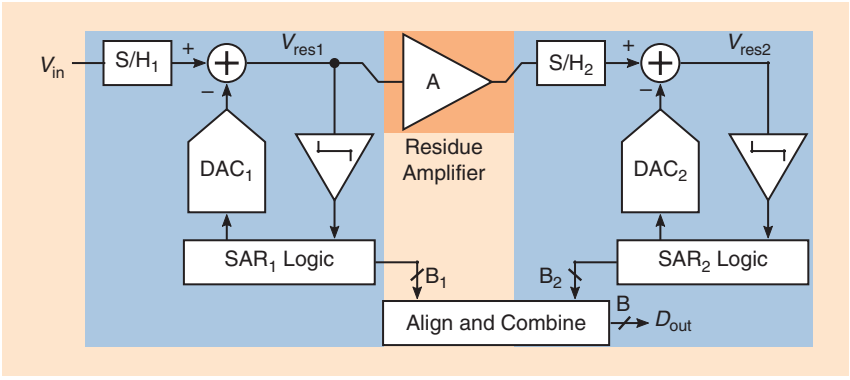


FIGURE 5: A *B*-bit 2-stage pipelined-SAR ADC.

found as in (5) by considering also the RA settling, resulting in

$$P_{P,\text{samp}} = V_{\text{DD}} \cdot NTC \cdot \vartheta_{P,\text{samp}} \cdot f_s \cdot \Big[\max\{C_S, C_{\min}\} + (2^{B_s}-1)\cdot \max\Big\{A_I C_I \cdot \Big(\frac{\vartheta_{P,\text{comp}} \cdot f_s}{\pi f_T}\Big), C_{\min}\Big\} + \max\Big\{A_s C_{\text{RA}} \cdot \Big(\frac{\vartheta_{P,\text{RA}} \cdot (B-(B_s-1)+2)\cdot \ln 2 \cdot f_s}{\zeta_{\text{set}} \cdot \pi f_T}\Big), C_{\min}\Big\}\Big] \cdot (\varphi_{\text{sup}} V_{\text{DD}}). \quad (19)$$

To derive $P_{P,\text{comp,tot}}$, the 2×-redundancy is considered for absorbing errors, thus relaxing the comparator power by scaling down C_I and C_L. An overrange relaxation factor ($\text{ORF} = 2^{B-B_s+1}$) is defined relating B_s, B, and $P_{P,\text{comp,tot}}$ with the contribution from all stages and both physical and process bounds converts to (20) at the bottom of the page.

Compared with (8) and (13), $(B - \log_2 A_I)$ is replaced with B_s in the latch part, which more correctly captures this contribution within the pipeline.

$P_{P,\text{lad}}$ is given by (9) as for the flash, with an extra *ORF* in the denominator. Finally, (10) gives the $P_{P,\text{dig}}$ of each sub-flash in the pipeline. On top, the "align and combine" logic is accounted for by adding the power of two extra gates/bit, leading to the final form

$$P_{P,\text{dig}} = V_{\text{DD}} \cdot f_s \cdot [5 \cdot (2^{B_s} - B_s) + 2 \cdot B_s] \cdot C_{\min} \cdot V_{\text{DD}}. \quad (21)$$

Pipelined-SAR Accuracy–Speed–Power Limits

The pipelined-SAR [15], [16] has been a blooming hybrid, appealingly combining its constituent architectures. A typical *B*-bit 2-stage pipelined-SAR (Figure 5) consists of a coarse B_1-bit SAR_1, a fine B_2-bit SAR_2, and an interstage RA.

The align and combine logic collects the bits and provides a new output with a latency of two clock periods at the speed of a single stage. Here, the RA can also be implemented in open loop and/or as integrator for potential power savings, whereas the S/H_2 can be integrated within the RA. Redundancy is incorporated in this hybrid as well, serving as a SAR_1 error absorption mechanism.

The combination of SAR and pipelining advances both its individual counterparts. With respect to a typical SAR, it boosts the converter speed by roughly the number of sequential cycles saved from each sub-SAR ($B_1, B_2 < B$). It also enables a higher efficiency when aiming for high accuracy by allowing lower-power sub-SAR comparators through the bit/stage allocation and residue amplification compared with the single-SAR comparator, whose power can get excessively high to meet the full accuracy requirements. With respect to a typical pipeline, it enhances the efficiency of multibit/stage approaches because of the linear power and parasitics increase of the sub-SAR as opposed to the exponential sub-flash. Therefore, it extends the range of resolutions, for which a multibit/stage approach is superior to its single-bit/stage counterpart, to values that allow a low-power sub-SAR comparator.

The majority of published pipelined-SAR ADCs are 2-stage. Few examples exist with a higher order of pipelining [17], [18]. Analogous to the regular pipeline, a higher-order pipelined-SAR intuitively achieves a higher absolute speed owing to sequential cycle reduction in each sub-SAR. A >2-stage pipelined-SAR should also show an increased efficiency when opting for higher-accuracy and gigahertz speed owing to reduced parasitic loading. This efficiency may be higher than the one of the regular pipeline provided that the sub-SAR is more efficient than the sub-flash for the same speed and resolution per stage. To investigate this, we begin by constructing a *B*-bit *M*-stage ($m = 2, 3, 4, 5$) converter with B_s-bit/stage and 2×-redundancy (B_s–1 effective bits). The bit/stage partitioning starts by first spreading *B* equally among the *m* stages, and the remaining bits (if any) are added sequentially from the last to the first stage (e.g., for $B = 11$ and $m = 4$, a 3-3-4-4 partitioning is followed). The generalized pipelined-SAR power consumption expression, including the sampler, the RA, the comparator, the CDAC, and the digital logic, is then

$$P_{PS,\text{tot}} = P_{PS,\text{samp}} + P_{PS,\text{RA,tot}} + B_s \cdot P_{PS,\text{comp,tot}} + P_{PS,\text{CDAC,tot}} + m \cdot P_{PS,\text{dig}}. \quad (22)$$

$$P_{P,\text{comp,tot}} = V_{DD} \cdot \vartheta_{P,\text{comp}} \cdot f_s \cdot \left\{ \frac{\left\{ \sum_{j=0}^{m-1} \max\left\{ \frac{C_I}{\text{ORF}} \cdot 2^{-(B_s-1)\cdot j}, C_{\min} \right\} \cdot \Delta V_I \right\}}{1 - \frac{2\Delta V_I}{\pi V_{GT}} \cdot \frac{\vartheta_{P,\text{comp}} \cdot f_s}{f_T}} \right\} + \left\{ \frac{[B_s \cdot \ln 2 + \ln \text{BER}^{-1}] \cdot \sum_{j=0}^{m-1} \max\left\{ \frac{C_L}{\text{ORF}} \cdot 2^{-(B_s-1)\cdot j}, C_{\min} \right\} \frac{V_{GT}}{2}}{1 - \frac{[B_s \cdot \ln 2 + \ln \text{BER}^{-1}]}{\pi} \cdot \frac{\vartheta_{P,\text{comp}} \cdot f_s}{f_T}} \right\}. \quad (20)$$

$P_{PS,\text{samp}}$ is derived including all the aforementioned laid-out assumptions and effects

$$P_{PS,\text{samp}} = V_{\text{DD}} \cdot NTC \cdot \vartheta_{PS,\text{samp}} \cdot f_s \cdot \left[\max\{C_S, C_{\min}\} + \max\left\{A_I C_I \cdot \left(\frac{\vartheta_{PS,\text{comp}} \cdot f_s}{\pi f_T}\right), C_{\min}\right\} + \max\left\{A_s C_{\text{RA}} \cdot \left(\frac{\vartheta_{PS,\text{RA}} \cdot (B-(B_s-1)+2) \cdot \ln 2 \cdot f_s}{\zeta_{\text{set}} \cdot \pi f_T}\right), C_{\min}\right\}\right] \cdot (\varphi_{\text{sup}} V_{\text{DD}}). \quad (23)$$

In our pipelined-SAR derivation, we allocate one and a half cycles to both the sampler and the RA, captured by $\vartheta_{PS,\text{samp}}$, $\vartheta_{PS,\text{RA}} = (B_s + 1 + 2)/1.5$. This is effectively seen as fitting one extra cycle in the total $1/f_s$-period, with $\vartheta_{PS,\text{comp}} = (B_s + 1 + 2) \cdot 1/\beta$.

$P_{PS,\text{RA,tot}}$ assumes the same $g_{m,\text{RA}}$-C_{RA} circuit with all the assumptions and effects already discussed as shown in (24) at the bottom of this page.

One distinct difference regarding the RA gain between the pipeline and pipelined-SAR is that in the former, A_s remains constant when increasing B since B_s is constant, whereas in the latter, A_s is a function of B_s since the bit partitioning is not the same when increasing B.

$P_{PS,\text{comp,tot}}$ is identical to (20) for $P_{P,\text{comp,tot}}$ by substituting $\vartheta_{P,\text{comp}}$ with $\vartheta_{PS,\text{comp}}$. For $P_{PS,\text{CDAC,tot}}$, the same switching as in (14) is considered with the appropriate sub-SAR switching activity, depending on the bit partitioning. Further, considering all stages under study, as well as the stage scaling, we get

$$P_{PS,\text{DAC,tot}} = \frac{V_{\text{DD}}}{\lambda_{\text{ref}}} \cdot \vartheta_{PS,\text{CDAC}} \cdot f_s \cdot \left[\sum_{i=2}^{5}\left(\sum_{j=1}^{B_s(i)-1} 2^{B_s(i)-3-2j} \cdot (2^j - 1)\right) \times \max\left\{\frac{C_S}{2^{B_s(i)}} \prod_{l=0}^{4} \frac{1}{A_{sl}(B_s(i))}, C_{\min}\right\}\right] \cdot V_{\text{REF}}. \quad (25)$$

TABLE 1. NOISE ALLOCATION PERCENTAGE (%) WITH RESPECT TO A CERTAIN TOTAL QUANTIZATION NOISE, $\overline{\varepsilon_q^2}$.

	SAMPLER	COMPARATOR	RA	LADDER
Full flash	50	35	—	15
Binary SAR	50	50	—	—
Pipeline	40	15	40	5
Pipelined-SAR	40	20	40	—

In the above, $l = 0$ corresponds to the CDAC of the first stage ($A_{sl} = 1$), while the CDAC timing portion over a full converter period is captured by $\vartheta_{PS,\text{CDAC}} = [(B_s + 1 + 2)/B_s] \cdot 1/(1-\beta)$. To conclude our derivations, $P_{PS,\text{dig}}$ assumes the same number of gates for each sub-SAR as for the regular SAR and adds two extra gates per bit in each for the align and combine logic

$$P_{PS,\text{dig}} = V_{\text{DD}} \cdot f_s \cdot (5 \cdot B_s + 2 \cdot B_s) \cdot C_{\min} \cdot V_{\text{DD}}. \quad (26)$$

Putting It All Together

To build insight on the discussed architectures' strengths and limitations, these are compared in terms of their accuracy–speed–power limits derived through (1)–(26) and plotted in Figure 6. Four deep-scaled CMOS processes are used, and the critical parameters are extracted from core low-threshold devices with interconnect. $C_{\min}$ is extracted as a 2×-minimum-sized inverter gate capacitance with equal PMOS/NMOS drivability. The allocated noise percentage to set the corresponding capacitors—C_S, C_I, C_L, and C_{RA}—are given in Table 1.

For low speed [$f_s = 500$ kHz; Figure 6(a)], the slopes of all the curves are first process-limited, with $C_{\min}$ setting the baseline, and then noise limited. Above about 40 dB SNDR, flash becomes the most energy inefficient, with a slope $\propto 2^B$. The pipelines show a similar slope when noise-limited but with a better overall efficiency owing to redundancy. The SAR is very hard to beat up to about 45 dB, after which its energy increases with a slope $\propto B$. The pipelined-SARs follow similar noise-limited slopes, showing the best efficiency between 45 and 65 dB while being slightly more efficient than the multibit/stage pipelines up to 70 dB. For medium-high speed [$f_s = 500$ MHz; Figure 6(b)], the flash remains the most inefficient above 40 dB. The others show steeper relative slopes owing to more stringent internal timings and increased parasitic effects through f_s/f_T. The SAR is still superior up to 40 dB. The fewer-stage pipelines and pipeline-SARs start losing the battle at high accuracies because the relaxed RA precision benefit is offset by the higher gain–bandwidth requirement. The 1,2-bit/stage pipelines and the 4,5-stage pipelined-SARs lead the efficiency at accuracies above 75 dB. For very high speed [$f_s = 1.3$ GHz; Figure 6(c)], the flash remains almost intact, while the exacerbated parasitic contribution in the others further deteriorates their efficiency. The SAR retains the lead up to 40 dB, while the 1,2-bit/stage pipelines lead above 75 dB. The 4,5-stage pipelined-SARs are winning between 40 and 65 dB and compete with the pipelines up to 75 dB. Extending the pipelined-SAR beyond five stages may even surpass the 1-bit/stage pipeline efficiency when targeting multi-gigahertz-range speed. For a similar stage count, the pipelined-SAR is potentially more efficient and faster than the pipeline for an extended range of accuracies. For each architecture, the derived limits are very similar across the

$$P_{PS,\text{RA,tot}} = A_s(B_s) \cdot \frac{\vartheta_{PS,\text{RA}}}{\zeta_{\text{set}}} \cdot f_s \cdot \frac{(B-(B_s-1)+2) \cdot \ln 2}{1 - A_s(B_s) \cdot \frac{(B-(B_s-1)+2) \cdot \ln 2}{\pi} \cdot \frac{\vartheta_{PS,\text{RA}} \cdot f_s}{\zeta_{\text{set}} \cdot f_T}} \times \left[\sum_{i=0}^{m-2} \max\{C_{\text{RA}} \cdot 2^{-(B_s-1)\cdot i}, C_{\min}\} + \sum_{i=0}^{m-2} \max\{C_{\text{in},I} \cdot 2^{-(B_s-1)\cdot i}, C_{\min}\}\right] \cdot \frac{V_{GT} \cdot V_{\text{DD}}}{2\eta_{\text{lin}}}. \quad (24)$$

different processes in the noise-limited region [Figure 6(d)]. The reason is that the analog scaling drawback of lowering V_{DD} is to a first-order nullified by the increased g_m/I_D to achieve the same f_T.

Although there are aspects omitted in our analysis (e.g., calibration, SAR redundancy, and asynchronous timing; unsettled integrator RA), the proposed "plug-and-play" framework captures to a first-order major fundamental, architecture, and process parameters influencing an ADC's efficiency. It makes a valuable toolbox when starting a new design or changing process nodes, offering a great insight into different circuit, architecture, and process capabilities. Further, it hints at a potentially superior ADC architecture for both high accuracy and high speed, the multistage (>5-stages) pipelined-SAR hybrid. Compared with the state-of-the-art design of Figure 1, it follows the relative accuracy and trends between the different architectures with an absolute accuracy of within an order of magnitude. This is fair, given the generalized nature of this framework. Finally, it greatly motivated the design choices of the state-of-the-art design in [18].

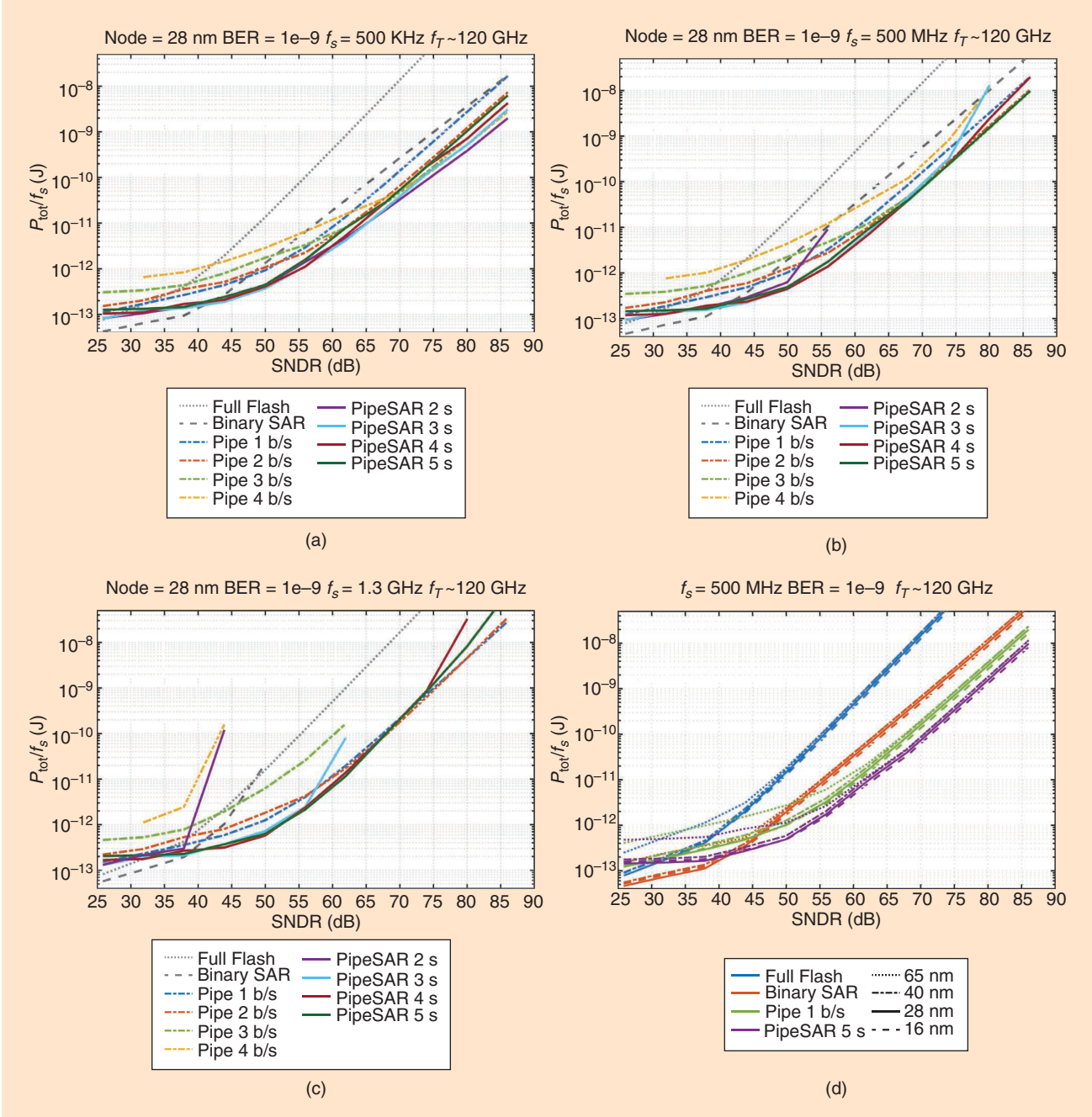


FIGURE 6: Derived architecture accuracy–speed–power limits: (a) f_s = 500 kHz, (b) f_s = 500 MHz, (c) f_s = 1.3 GHz, and (d) across different processes.

References

[1] B. Murmann, "ADC Performance Survey 1997-2021." Accessed: Aug. 2021. [Online]. Available: https://web.stanford.edu/~murmann/adcsurvey.html

[2] B. Murmann, "Limits on ADC power dissipation," in *Analog Circuit Design*, M. Steyaert, J. Huijsing, and A. van Roermund Eds. Berlin, Germany: Springer-Verlag, 2006.

[3] T. Sundström, B. Murmann, and C. Svensson, "Power dissipation bounds for high-speed nyquist analog-to-digital converters," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, no. 3, pp. 509–518, Mar. 2009, doi: 10.1109/TCSI.2008.2002548.

[4] A. Varzaghani *et al.*, "A 10.3-GS/s, 6-bit Flash ADC for 10G Ethernet Applications," *IEEE J. Solid-State Circuits*, vol. 48, no. 12, pp. 3038–3048, Dec. 2013, doi: 10.1109/JSSC.2013.2279419.

[5] V. H.-C. Chen and L. Pileggi, "A 69.5 mW 20 GS/s 6b time-interleaved ADC with embedded time-to-digital calibration in 32 nm CMOS SOI," *IEEE J. Solid-State Circuits*, vol. 49, no. 12, pp. 2891–2901, Dec. 2014, doi: 10.1109/JSSC.2014.2364043.

[6] W. Sansen, "Analog CMOS from 5 micrometer to 5 nanometer," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC)*, San Francisco, CA, USA, Feb. 2015, pp. 1–6, doi: 10.1109/ISSCC.2015.7062848.

[7] M. J. M. Pelgrom, A. C. Duinmaijer, and A. P. Welbers, "Matching properties of MOS transistors," *IEEE J. Solid-State Circuits*, vol. 24, no. 5, pp. 1433–1439, 1989, doi: 10.1109/JSSC.1989.572629.

[8] M. J. Kramer, E. Janssen, K. Doris, and B. Murmann, "A 14b 35 MS/s SAR ADC achieving 75 dB SNDR and 99 dB SFDR with loop-embedded input buffer in 40 nm CMOS," *IEEE J. Solid-State Circuits*, vol. 50, no. 12, pp. 2891–2900, Dec. 2015, doi: 10.1109/JSSC.2015.2463110.

[9] A. T. Ramkaj, M. Strackx, M. S. J. Steyaert, and F. Tavernier, "A 1.25-GS/s 7b SAR ADC with 36.4-dB SNDR at 5 GHz using switch-bootstrapping, USPC DAC and triple-tail comparator in 28-nm CMOS," *IEEE J. Solid-State Circuits*, vol. 53, no. 7, pp. 1889–1901, Jul. 2018, doi: 10.1109/JSSC.2018.2822823.

[10] V. Hariprasath, J. Guerber, S.-H. Lee, and U.-K. Moon, "Merged capacitor switching based SAR ADC with highest switching energy-efficiency," *Electron. Lett.*, vol. 46, no. 9, pp. 620–621, May 2010, doi: 10.1049/el.2010.0706.

[11] 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, doi: 10.1109/JSSC.2017.2747758.

[12] 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, doi: 10.1109/JSSC.2020.3023882.

[13] Y. Chiu, P. R. Gray, and B. Nikolic, "A 14-b 12-MS/s CMOS pipeline ADC with over 100-dB SFDR," *IEEE J. Solid-State Circuits*, vol. 39, no. 12, pp. 2139–2151, Dec. 2004, doi: 10.1109/JSSC.2004.836232.

[14] B. Verbruggen *et al.*, "A 2.6mW 6b 2.2GS/s 4-times Interleaved fully dynamic pipelined ADC in 40nm digital CMOS," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC)*, San Francisco, CA, USA, Feb. 2010, pp. 296–297, doi: 10.1109/ISSCC.2010.5433925.

[15] F. van der Goes *et al.*, "A 1.5 mW 68 dB SNDR 80 MS/s 2x Interleaved Pipelined SAR ADC in 28 nm CMOS," *IEEE J. Solid-State Circuits*, vol. 49, no. 12, pp. 2835–2845, Dec. 2014, doi: 10.1109/JSSC.2014.2361774.

[16] 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, doi: 10.1109/JSSC.2015.2464684.

[17] B. Vaz *et al.*, "A 13b 5GS/s ADC with time-interleaved chopping calibration in 16nm FinFET," in *Proc. IEEE Symp. VLSI Circuits (VLSI Circuits)*, Honolulu, Hawaii, Jun. 2018, pp. C99–C100, doi: 10.1109/VLSIC.2018.8502306.

[18] A. T. 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, doi: 10.1109/JSSC.2019.2960476.

About the Authors

Athanasios Ramkaj (aramkaj@stanford.edu) received his M.Sc. degree (*cum laude*) in electrical engineering from the Technische Universiteit Delft, The Netherlands, and his Ph.D. degree (*summa cum laude*) in electrical engineering from the Katholieke Universiteit Leuven, Belgium, in 2014 and 2021, respectively. He is currently a postdoctoral researcher with Stanford University, Stanford, California, 94305, USA. From 2013 to 2014, he was an intern with NXP Semiconductors, Eindhoven, The Netherlands, investigating gigahertz-range analog-to-digital converters (ADCs). In 2019, he interned with Analog Devices, Wilmington, Massachusetts, USA, researching RF ADC bandwidth extension solutions. He received the 2021 Analog Devices Outstanding Student Designer Award and the 2020 IEEE SSCS Predoctoral Achievement Award. His main research interests include high-speed/bandwidth high-resolution RF-sampling ADCs, high-speed analog/mixed-signal circuits, and ultra-wideband front ends.

Marcel J.M. Pelgrom received his M.Sc. and Ph.D. degrees from Twente University, The Netherlands. In 1979 he joined Philips Research, where he headed the high-speed analog-to-digital conversion team and served seven years as the department head for mixed-signal electronics research. His research covered metal-oxide semiconductor mismatch, basic circuits, and analog-to-digital converters. The most cited article in the *IEEE Solid-State Circuits Magazine* is his article on metal-oxide semiconductor matching properties. Until 2013, he was a member of the technical staff of NXP Semiconductors, involved in the research on variability, and advanced conversion techniques. In 2003 and 2014, he spent a sabbatical at Stanford University, Stanford, California, USA. He was the 2017 IEEE Gustav R. Kirchhoff Award recipient. At present, he is an independent consultant.

Michiel S.J. Steyaert received his M.Sc. degree in electrical and mechanical engineering and his Ph.D. degree in electronics from the Katholieke Universiteit, Belgium, in 1983 and 1987, respectively. In 1988, he was a visiting professor with the University of California, Los Angeles, California, USA. He was a research associate in 1989, a senior research associate in 1992, and a research director in 1996, with the Department of Electrical Engineering, ESAT. From 2005 until 2012, he was the Electrical Engineering Department chair, where he is currently a full professor. His research lies with high-performance, high-frequency analog circuits for telecommunication systems and analog signal processing. He is the only European researcher who has received both the 50th and the 60th Anniversary Top ISSCC Contributors Award in 2003 and 2013.

Filip Tavernier (filip.tavernier@esat.kuleuven.be) received his M.Sc. degree in electrical engineering and his Ph.D. degree in engineering science from the Katholieke Universiteit (KU Leuven), Belgium, in 2005 and 2011, respectively. From 2011 to 2014, he was a Senior Fellow in the Microelectronics Group at CERN, Geneva, Switzerland. From 2014 to 2015, he was a postdoctoral researcher with the Department of Electrical Engineering, KU Leuven. From 2015 until 2020, he was an assistant professor, and since 2020, he has been an associate professor at KU Leuven, Leuven, Belgium, within the same department. His main research interests are analog and mixed-signal integrated circuits for high-performance data converters, dc–dc converters, optical receivers, and cryogenic circuits.